Department of AERONAUTICS and ASTRONAUTICS
STANFORD UNIVERSITY

68 p, Cal.

8299000

[redacted]

G. W. DELEY

UNPUBLISHED PRELIMINARY DATA

OPTIMAL BOUNDED CONTROL OF LINEAR SAMPLED-DATA SYSTEMS USING QUADRATIC PERFORMANCE CRITERIA

THIS WORK WAS PERFORMED IN ASSOCIATION WITH RESEARCH SPONSORED BY
THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
UNDER RESEARCH GRANT NsG-133-61

APRIL 1963

SUDAER NO. 148

Department of Aeronautics and Astronautics
Stanford University
Stanford, California

OPTIMAL BOUNDED CONTROL OF LINEAR SAMPLED-DATA SYSTEMS USING QUADRATIC PERFORMANCE CRITERIA

by

Gary W. Deley

SUDAER No. 148
April 1963

This work was performed in association with research sponsored by the National Aeronautics and Space Administration under Research Grant NsG-133-61

ABSTRACT

14290

This investigation studies optimal control of linear sampled-data systems where the control is subject to saturation. The system is described by the state-space method. The control is considered to be optimal when it minimizes a performance index which is defined as a sum over the sampling instants of a quadratic function of the states and controls.

The solution begins with the Principle of Optimality. A form is assumed for the optimal return function, and recurrence relations are derived for the one-input case which are different depending on whether the optimal control is or is not saturated. The optimal control is shown to be a piecewise linear function of the states. A computing method that uses the recurrence relations to solve the infinite stage regulator problem is presented and discussed in detail. This method requires less computer time and memory than would straight dynamic programming.

Both one- and two-input control are considered. The two-input case requires a third set of recurrence relations for use when one input is saturated and the other is not. More inputs can be handled using the same methods, but the complexity increases rapidly with the number of inputs. A detailed discussion of a simple method for finding the minimum of a positive definite quadratic function in two variables subject to the constraint that the minimum be on or within a rectangle is presented.

Four examples showing the optimal control of second-order systems determined by the computing method given in this report are presented and discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

LIST OF SYMBOLS

a_i	an element of A_{N+1}, scalar control case
b	the element of B_{N+1}, scalar control case
k	a constant of proportionality
k_{ij}	an element of the matrix K
n	an integer
s	Laplace transform argument
t	continuous time variable
u	a scalar control
$\underline{u}$	a control vector
u'	a scalar
$\underline{u}'$	a vector
u*	control input to a hold
u_i	an element of $\underline{u}$
u_i'	an element of $\underline{u}'$
u_i''	a scalar
$\underline{y}$	an output vector
$\underline{y}_d$	desired output vector
$\underline{z}$	a state vector
z_i	an element of $\underline{z}$
A_{N+1}	a feedback coefficient matrix
$A'_{i\,N+1}$	a feedback coefficient vector
B_{N+1}	a constant feedback vector
$B'_{i\,N+1}$	a feedback constant
C_N	a scalar
I	moment of inertia

LIST OF SYMBOLS (Cont)

I_N	optimal return function for N stages
J_N	a performance index
K	a positive definite symmetric matrix
M	an output matrix
N	an integer
P_N	a positive semidefinite symmetric matrix
Q	a positive semidefinite symmetric matrix
Q'	a positive semidefinite symmetric matrix
R_N	a vector
T	(superscript) transpose operator
U(s)	Laplace transform of u
Y_n	a performance number
$Z_i(s)$	Laplace transform of z_i
α	represents either α^+ or α^-
α_i	represents an element of either $\underline{\alpha}^+$ or $\underline{\alpha}^-$
α_i^-	an element of $\underline{\alpha}^-$
α_i^+	an element of $\underline{\alpha}^+$
α^+, α^-	upper and lower bounds on the scalar control
$\underline{\alpha}^+$, $\underline{\alpha}^-$	upper and lower bounds on the control vector
β	a constant
γ_{ij}	an element of the matrix Γ
δ_{ij}	an element of the matrix Δ
θ	attitude error angle
ξ	a positive definite function
τ	sampling interval

LIST OF SYMBOLS (Cont)

ϕ_{ij}	an element of the matrix Φ
Γ	a positive semidefinite symmetric matrix
Δ	a distribution matrix
Δ_i	a submatrix partitioned from Δ
Φ	a transition matrix

ACKNOWLEDGMENT

Appreciation is expressed for the guidance of Dr. Gene F. Franklin, under whom this research was conducted, and for the helpful suggestions of Dr. Robert H. Cannon, Jr. in the presentation of the material.

I. INTRODUCTION

A. OUTLINE OF THE PROBLEM

As an example of the problem investigated in this report, consider a space vehicle whose attitude is to be controlled by reaction wheels or gyros. Various disturbances, perhaps impulsive as from collisions with micrometeorites, produce an error in attitude which must be corrected. In applying the control, the integral over time of the attitude squared-error plus the squared-control is to be minimized.

Because the control torque is subject to saturation, a nonlinearity is inherent in the system. Thus it is expected that the optimal control will be a nonlinear function of the states of the system--the attitude error and velocity. This optimal control function is to be stored in a small special-purpose digital computer called a digital controller.

The introduction of a computer makes the system sampled-data. The computer determines from the states of the system at each sampling instant the correct optimal control to apply over the next sampling interval.

Since the system is now sampled-data, rather than minimize an integral it is logical and convenient to minimize the sum over time of the attitude squared-error plus the squared-control at the sampling instants.

More generally, this investigation studies optimal control of linear sampled-data systems where the control is subject to saturation. The system will be described by the state-space method developed by Kalman and Bertram [Ref. 1].

At each sampling instant the system is assigned a performance number, which is a quadratic function of the state error and control. The sum of the performance numbers over a given number of samples is called the performance index. Only the transient regulator problem--that of finding the control sequence which, from a given initial condition with no external disturbances and no commands, minimizes the performance index--will be considered.

The sampling rate is often fast enough that a sampled-data system can be closely approximated, for the purpose of analysis, by a continuous system. In this investigation, however, the sampling rate is considered

to be slow enough that the sampling process introduces significant effects into the performance of the system.

B. SUMMARY OF RELATED WORK

There is a considerable body of literature on the subject of optimal control of sampled-data systems, but almost no mention of the specific problem presented here.

Using the Principle of Optimality, the problem without constraints on the control was solved by Kalman and Koepcke [Ref. 2]. They show that for the infinite stage regulator problem the optimal control takes the form of stationary, linear feedback gains. Work on this problem was also done by Henry [Ref. 3].

Several researchers have worked on the problem investigated here, though using minimum time response as the criterion of optimality. Among these are Kurzweil [Ref. 4], Desoer and Wing [Ref. 5], and Kalman [Ref. 6].

Merrian [Refs. 7, 8], using his parametric expansion method, has studied the problem in the continuous case.

Bellman's computational method of dynamic programming [Ref. 9] solves, among others, problems of the type studied here when the dimension of the state vector is small. The special problem of this report, minus constraints on the control is mentioned by Bellman and Dreyfus [Ref. 10].

Quadratic performance criteria have been used by many researchers in both the continuous and sampled-data cases.

The state-space method of describing linear sampled-data systems is discussed in detail by Kalman and Bertram [Ref. 1], Kalman [Ref. 11], Gunckel [Ref. 12], and Rauch [Ref. 13].

C. OUTLINE OF NEW RESULTS

For the first time in the literature the problem of computing the optimal feedback coefficients of a sampled-data system with bounded control using quadratic performance criteria is discussed in detail.

In Chapter II a mathematical description of the system and the performance criterion is given, and the problem formulation is presented. Two examples using this formulation are discussed.

Recurrence relations necessary to the computation method are derived in Chapter III for the single-control case. The optimal control is shown to be a piecewise-linear function of the states.

In Chapter IV a general computing method is presented for the single-input case, and problems connected with the computations are discussed in detail. The method is also compared with dynamic programming. It is shown that, because it takes advantage of the information contained in the recurrence relations, the method developed here requires much less computer time and memory than would dynamic programming.

Chapter V extends the work to the case where the system has two controlling inputs. Extension to systems with more inputs presents no formal difficulties but is not discussed due to its complexity.

Results of computer solutions of four examples are presented and discussed in Chapter VI.

II. STATEMENT OF THE PROBLEM

This investigation considers those sampled-data systems that can be adequately described by linear-difference equations. These equations will be written in the state-space form used by Kalman and others. For conciseness, vector-matrix notation will be used throughout.

A. THE SYSTEM

The plant, or system to be controlled, is described by the linear vector-difference equation

$$\underline{z}(n+1) = \Phi\underline{z}(n) + \Delta\underline{u}(n) \tag{2.1}$$

and the vector equation

$$\underline{y}(n) = M\underline{z}(n), \tag{2.2}$$

where $\underline{z}(n)$ is an $(m \times 1)$ state vector,
$\underline{y}(n)$ is a $(p \times 1)$ output vector,
$\underline{u}(n)$ is a $(q \times 1)$ input (control) vector,
Φ is an $(m \times m)$ transition matrix,
Δ is an $(m \times q)$ distribution matrix,
M is a $(p \times m)$ output matrix.

All vectors are considered to be column vectors. Row vectors will be written, for example, as $\underline{z}^T(n)$, where T denotes the transpose operation. $\underline{y}(n)$ is the measurable output vector. If all the states are directly measurable, then M is the identity matrix.

Since physically the control variables cannot be unbounded, each element of the control vector $\underline{u}(n)$ is bounded from below by the corresponding element of a vector $\underline{\alpha}^-$ and from above by the vector $\underline{\alpha}^+$. That is,

$$\underline{\alpha}^- \leq \underline{u}(n) \leq \underline{\alpha}^+. \tag{2.3}$$

The control vector $\underline{u}(n)$ will have dimension one (i.e., it will be a scalar) in Chapters III and IV. Chapter V will extend the results to higher dimensional $\underline{u}(n)$.

B. THE PERFORMANCE CRITERION

If a system is to be optimized, some criterion must be chosen that determines how well the system is operating. In this investigation a single number that characterizes overall performance is assigned to the system at each sampling instant. This number, called the performance number Y_n, is defined to be a quadratic function of the difference between the actual output of the system, $\underline{y}(n)$, and the constant desired output $\underline{y}_d$ plus a quadratic cost on the control required to achieve the output. Mathematically this is

$$Y_n = [\underline{y}(n) - \underline{y}_d]^T Q' [\underline{y}(n) - \underline{y}_d] + \underline{u}^T(n-1)\Gamma\underline{u}(n-1), \qquad (2.4)$$

where Q' and Γ are positive semidefinite symmetric matrices. With no further loss in generality let $\underline{y}_d = 0$.

Y_n can be stated in terms of $\underline{z}(n)$ by using Eq. (2.2).

$$Y_n = \underline{z}^T(n)Q\underline{z}(n) + \underline{u}^T(n-1)\Gamma\underline{u}(n-1), \qquad (2.5)$$

where Q is a symmetric positive semidefinite matrix defined by

$$Q = M^T Q' M. \qquad (2.6)$$

Given an initial condition $\underline{z}(0)$, the control is considered optimal if it minimizes in N stages the sum of the costs Y_n. This sum, called the performance index, is denoted by $J_N[\underline{z}(0)]$.

$$J_N[\underline{z}(0)] = \sum_{n=1}^{N} [\underline{z}^T(n)Q\underline{z}(n) + \underline{u}^T(n-1)\Gamma\underline{u}(n-1)] \qquad (2.7)$$

Although the performance index is limited to quadratic functions, many useful problems can be formulated using criteria of this type. Integral-squared-error has been used with continuous systems for some time, and sum-squared-error is a logical extension to use with sampled-data systems. The above formulation allows not only squared-error terms, but also cross-products between the states, to be charged. Often

the energy used for control must be conserved, and the charge on squared-control allows for this. Squared terms also provide a simple analytical approximation to absolute value.

The principal concern of this investigation is the infinite stage regulator problem; thus the performance index is

$$J_{\infty}[\underline{z}(0)] = \lim_{N \to \infty} J_N[\underline{z}(0)]. \tag{2.8}$$

C. THE PROBLEM STATEMENT

The problem can now be precisely stated: Given the system defined by the linear vector-difference Eq. (2.1), and given the bounds on the control defined by the vector-inequality (2.3), find for all initial conditions $\underline{z}(0)$ the control sequences $\underline{u}[\underline{z}(0)]$, $\underline{u}[\underline{z}(1)]$, $\underline{u}[\underline{z}(2)]$, ... that minimize the performance index $J_{\infty}[\underline{z}(0)]$.

Finding the optimal control for all states distinguishes the control problem from the optimal trajectory problem. In the latter usually only one or a few initial states are of interest.

D. EXAMPLES

Two examples of the above formulation will be given. The solution to these examples will be discussed in Chapter VI.

For the first example consider a space vehicle whose attitude is to be controlled to an inertially fixed reference direction by reaction wheels. In its simplest formulation the small angular motion of the vehicle about a principal axis can be studied by considering the vehicle as an inertia with moment of inertia I about that axis [Ref. 14]. The state variables are the attitude error θ and its derivative $\dot{\theta} = d\theta/dt$. The sampling interval is τ seconds long, and the control is held constant over the sampling interval by a zero-order hold [Ref. 15]. The system is shown in Fig. 1.

The equations of motion are

$$\dot{z}_1(t) = z_2(t)$$

$$\dot{z}_2(t) = \frac{u(t)}{I}. \tag{2.9}$$

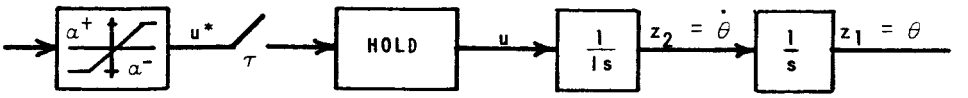

FIG. 1. BLOCK DIAGRAM OF SYSTEM IN FIRST EXAMPLE.

The hold takes the value of u^* at the sampling instant and holds it constant at that value until the next sampling instant. That is,

$$u(t) = u^*(n\tau), \qquad \text{for} \quad n\tau \leq t < (n+1)\tau. \tag{2.10}$$

Solving Eqs. (2.9) for $z_1[(n+1)\tau]$ and $z_2[(n+1)\tau]$ in terms of $z_1(n\tau)$, $z_2(n\tau)$, and $u(n\tau)$ gives the Φ and Δ matrices. A simple way to determine these matrices is to let, one at a time, an independent variable z_1, z_2, or u at time $n\tau$ be unity while the others are zero and solve for the dependent variables z_1 and z_2 at time $(n+1)\tau$. Thus, for example, let the Laplace transform of $u(t - n\tau)$ be $U(s) = 1/s$ and solve for $Z_1(s)$, which is

$$Z_1(s) = \frac{1}{Is^2}\,U(s) = \frac{1}{Is^3}. \tag{2.11}$$

The inverse transform is

$$z_1(t - n\tau) = \frac{(t - n\tau)^2}{2I}. \tag{2.12}$$

Letting $t = (n+1)\tau$ gives $\delta_{11}(\tau)$.

$$\delta_{11}(\tau) = \frac{\tau^2}{2I}. \tag{2.13}$$

In the same manner the other elements of the Φ and Δ matrices can be found. These are

$$\Phi = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix}, \qquad \Delta = \frac{1}{I}\begin{bmatrix} \tau^2/2 \\ \tau \end{bmatrix}. \tag{2.14}$$

A performance criterion needs to be chosen. M is the identity matrix, which means both attitude error and its rate of change can be measured directly. Assume the performance number is the sum of the attitude squared-error and the squared-control. Furthermore, assume the cost of an error in attitude is to be weighted equally with the cost of control. Thus

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \Gamma = 1. \tag{2.15}$$

The problem then is: For each initial condition $\underline{z}(0)$ find the control sequence $u(0)$, $u(1)$, ... that minimizes the performance index

$$J_{\infty}[\underline{z}(0)] = \sum_{n=1}^{\infty} \{z_1^2(n\tau) + u^2[(n-1)\tau]\}. \tag{2.16}$$

From here on, to conform with the original problem statement, the τ will be dropped from the arguments, with no implication that $\tau = 1$.

As a second example consider an artificial satellite orbiting the earth. Using small angle approximations and neglecting other terms of small magnitude, the pitch equations of motion are decoupled from roll and yaw. The vehicle can be described in pitch as an inertia with moment of inertia I [Refs. 14, 16]. An important external force acting on the satellite is exerted by the gravity gradient. For small values of θ, this force is proportional to the attitude error θ with constant of proportionality k, as shown in Fig. 2.

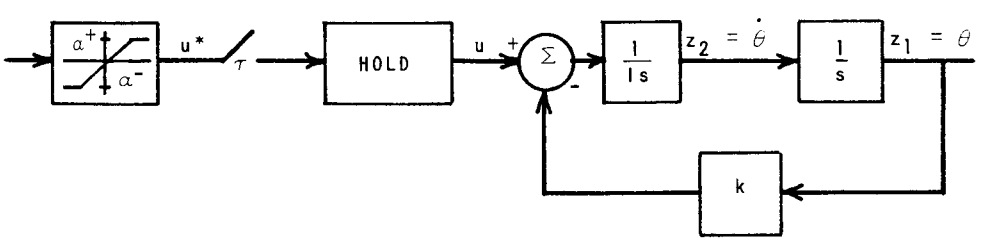

FIG. 2. BLOCK DIAGRAM OF SYSTEM IN SECOND EXAMPLE.

From Fig. 2 the equations of motion can be written down by inspection.

$$\dot{z}_1(t) = z_2(t)$$

$$\dot{z}_2(t) = -(k/I)z_1(t) + u(t)/I \tag{2.17}$$

For this example the method of finding ϕ_{11} will be shown in detail. Conceptually it is easier in this case to consider the transfer function from $Z_2(s)$ to $Z_1(s)$ and let $z_2(n)$ be the delta function. This has the effect of making $z_1(n) = 1$ as desired. Thus

$$Z_1(s) = \frac{s}{s^2 + (k/I)^2} Z_2(s); \qquad Z_2(s) = 1. \tag{2.18}$$

Therefore

$$\phi_{11}(\tau) = \cos(\beta\tau) \tag{2.19}$$

where

$$\beta^2 = k/I. \tag{2.20}$$

Similarly the entire Φ and Δ matrices can be found, and the vector-difference equation is written as

$$\underline{z}(n+1) = \begin{bmatrix} \cos(\beta\tau) & (1/\beta)\sin(\beta\tau) \\ -\beta\sin(\beta\tau) & \cos(\beta\tau) \end{bmatrix} \underline{z}(n) + \frac{1}{I} \begin{bmatrix} (1/\beta^2)[1 - \cos(\beta\tau)] \\ (1/\beta)\sin(\beta\tau) \end{bmatrix} u(n). \tag{2.21}$$

The performance index for this example is chosen as

$$J_\infty[\underline{z}(0)] = \sum_{n=1}^{\infty} z_1^2(n). \tag{2.22}$$

The solutions to both of the preceding examples are discussed in detail in Chapter VI.

III. RECURRENCE RELATIONS

A. SOLUTION WITH UNBOUNDED CONTROL

Before considering the case where the control $\underline{u}(n)$ is bounded, the solution to the unbounded control problem will be derived in detail. Here there is no simplification in having $\underline{u}(n)$ a scalar. The system is

$$\underline{z}(n+1) = \Phi\underline{z}(n) + \Delta\underline{u}(n) \tag{3.1}$$

$$\underline{y}(n) = M\underline{z}(n). \tag{3.2}$$

Given an initial condition $\underline{z}(0)$ the control sequence $\underline{u}(0)$, $\underline{u}(1)$, ..., $\underline{u}(N-1)$ is to be found that minimizes the performance index

$$J_N[\underline{z}(0)] = \sum_{n=1}^{N} [\underline{z}^T(n)Q\underline{z}(n) + \underline{u}^T(n-1)\Gamma\underline{u}(n-1)]. \tag{3.3}$$

The solution begins by defining $I_N[\underline{z}(0)]$, called the optimal return function, as the minimum value of $J_N[\underline{z}(0)]$. This $I_N[\underline{z}(0)]$ has a known and simple form:

$$I_N[\underline{z}(0)] = \underline{z}^T(0)P_N\underline{z}(0), \tag{3.4}$$

where P_N is a symmetric, positive semidefinite matrix. That this form is correct will be proved by induction later.

By definition

$$I_{N+1}[\underline{z}(0)] = \min_{\underline{u}(0)} \min_{\underline{u}(1)} \cdots \min_{\underline{u}(N)} \sum_{n=1}^{N+1} [\underline{z}^T(n)Q\underline{z}(n) + \underline{u}^T(n-1)\Gamma\underline{u}(n-1)]. \tag{3.5}$$

Since $\underline{z}(1)$ is determined solely by the choice of $\underline{u}(0)$ and not by the other $\underline{u}(n)$, Eq. (3.5) can be factored as

$$I_{N+1}[\underline{z}(0)] = \min_{\underline{u}(0)} \left\{ \underline{z}^T(1)Q\underline{z}(1) + \underline{u}^T(0)\Gamma\underline{u}(0) + \min_{\underline{u}(1)} \min_{\underline{u}(2)} \cdots \min_{\underline{u}(N)} \sum_{n=2}^{N+1} [\underline{z}^T(n)Q\underline{z}(n) + \underline{u}^T(n-1)\Gamma\underline{u}(n-1)] \right\} . \tag{3.6}$$

Noting that the last group of terms is exactly $I_N[\underline{z}(1)]$ gives

$$I_{N+1}[\underline{z}(0)] = \min_{\underline{u}(0)} \{\underline{z}^T(1)Q\underline{z}(1) + \underline{u}^T(0)\Gamma\underline{u}(0) + I_N[\underline{z}(1)]\} . \tag{3.7}$$

The above equation could have been arrived at directly using the Principle of Optimality [Ref. 9]. This principle states that the minimum cost of an N+1 stage process is the minimum of the sum of the cost of the first stage and the minimum cost of the remaining N stages. (Note that the arguments of the state and control variables increase with time, while the subscript on I_N decreases with time.)

Substituting Eqs. (3.1) and (3.4) into (3.7) gives

$$I_{N+1}[\underline{z}(0)] = \min_{\underline{u}(0)} \{[\Phi\underline{z}(0) + \Delta\underline{u}(0)]^T(Q + P_N)[\Phi\underline{z}(0) + \Delta\underline{u}(0)] + \underline{u}^T(0)\Gamma\underline{u}(0)\} . \tag{3.8}$$

Completing the square on the right side of (3.8) and defining

$$\underline{u}'(0) = A_{N+1}\underline{z}(0) \tag{3.9}$$

$$A_{N+1} = -[\Delta^T(Q + P_N)\Delta + \Gamma]^{-1}\Delta^T(Q + P_N)\Phi \tag{3.10}$$

transforms Eq. (3.8) into Eq. (3.11):

$$I_{N+1}[\underline{z}(0)] = \min_{\underline{u}(0)} \{[\underline{u}(0) - \underline{u}'(0)]^T[\Delta^T(Q + P_N)\Delta + \Gamma][\underline{u}(0) - \underline{u}'(0)] + \underline{z}^T(0)\Phi^T(Q + P_N)\Phi\underline{z}(0) - \underline{z}^T(0)A_{N+1}^T[\Delta^T(Q + P_N)\Delta + \Gamma]A_{N+1}\underline{z}(0)\} . \tag{3.11}$$

The control $\underline{u}(0)$ occurs in only the first term of (3.11). If the matrix $[\Delta^T(Q + P_N)\Delta + \Gamma]$ is positive definite the optimal control is unique. Then the minimum value, zero, of this first term occurs only at

$$\underline{u}(0) = \underline{u}'(0). \tag{3.12}$$

The matrix will be positive definite if Γ is positive definite or if Q is positive definite and the columns of Δ are linearly independent. It will not be positive definite if $\Gamma = 0$ and the columns of Δ are linearly dependent [Ref. 12]. In other cases this matrix might be singular, although no such difficulty was encountered in the examples of Chapter VI.

Equation (3.12), along with Eqs. (3.9) and (3.10), defines the optimal value of $\underline{u}(0)$.

The recurrence relation for P_{N+1} is determined by equating Eq. (3.11) with (3.4) when $\underline{u}(0) = \underline{u}'(0)$.

$$\begin{aligned} I_{N+1}[\underline{z}(0)] &= \underline{z}^T(0)\Phi^T(Q + P_N)\Phi\underline{z}(0) \\ &\quad - \underline{z}^T(0)A_{N+1}^T[\Delta^T(Q + P_N)\Delta + \Gamma]A_{N+1}\underline{z}(0) \\ &= \underline{z}^T(0)P_{N+1}\underline{z}(0). \end{aligned} \tag{3.13}$$

Since (3.13) must hold for all $\underline{z}(0)$ the recurrence relation becomes

$$P_{N+1} = \Phi^T(Q + P_N)(\Phi + \Delta A_{N+1}), \tag{3.14}$$

where the relation

$$-A_{N+1}^T[\Delta^T(Q + P_N)\Delta + \Gamma] = \Phi^T(Q + P_N)\Delta \tag{3.15}$$

has been used to simplify (3.14).

Equation (3.13) shows that if the quadratic form for I_N is correct, then I_{N+1} has the same form. The quadratic form is trivially correct for I_0 since

$$I_0[\underline{z}(0)] = 0 \tag{3.16}$$

for all $\underline{z}(0)$. To complete the mathematical induction the form for I_1 must be shown to be correct. I_1 is determined from Eqs. (3.13) and (3.10) noting that $P_0 = 0$. Equation (3.15) is again used to simplify the result.

$$I_1[\underline{z}(0)] = \underline{z}^T(0)\Phi^T Q[\Phi - \Delta(\Delta^T Q\Delta + \Gamma)^{-1}\Delta^T Q\Phi]\underline{z}(0). \tag{3.17}$$

Thus $I_1[\underline{z}(0)]$ has the required quadratic form.

The solution proceeds as follows: Since $I_0[\underline{z}(0)] = 0$, $P_0 = 0$. Beginning with $P_0 = 0$ calculate A_1. From A_1 and P_0 calculate P_1. This iteration process is continued until all the A_N of interest are calculated. If the plant is controllable the A_N will tend to a limit as N increases [Refs. 11, 12]. Therefore, for the infinite stage regulator problem the optimal control in the unbounded case takes the form of a stationary, linear function of the states.

B. RECURRENCE RELATIONS WITH BOUNDED CONTROL

In the first part of this section the control $\underline{u}(n)$ will be a vector of any dimension. This will make it possible to use Eqs. (3.18) through (3.25) in Chapter V, where two-dimensional control is considered in detail. When the actual minimization over $u(0)$ is done in this section, $u(0)$ will be considered a scalar.

Limiting the possible range of the control $\underline{u}(n)$ to

$$\underline{\alpha}^- \leq \underline{u}(n) \leq \underline{\alpha}^+ \tag{3.18}$$

complicates the solution greatly. The derivation in this section is the same as that in Sec. A up to Eq. (3.7). Equation (3.7) becomes

$$I_{N+1}[\underline{z}(0)] = \min_{\underline{\alpha}^- \leq \underline{u}(0) \leq \underline{\alpha}^+} \{\underline{z}^T(1)Q\underline{z}(1) + \underline{u}^T(0)\Gamma\underline{u}(0) + I_N[\underline{z}(1)]\}. \tag{3.19}$$

$I_N[\underline{z}(0)]$ takes the form, as will later be proved by induction,

$$I_N[\underline{z}(0)] = \underline{z}^T(0)P_N\underline{z}(0) + \underline{z}^T(0)R_N + R_N^T\underline{z}(0) + C_N \tag{3.20}$$

where

P_N is an (m x m) positive semidefinite symmetric matrix,

R_N is an (m x 1) vector,

C_N is a scalar.

Substituting (3.1) and (3.20) into (3.19) gives

$$I_{N+1}[z(0)] = \min_{\underline{\alpha}^- \le \underline{u}(0) \le \underline{\alpha}^+} \{ [\Phi \underline{z}(0) + \Delta \underline{u}(0)]^T (Q + P_N)[\Phi \underline{z}(0) + \Delta \underline{u}(0)] + \underline{u}^T(0) \Gamma \underline{u}(0) + [\Phi \underline{z}(0) + \Delta \underline{u}(0)]^T R_N + R_N^T [\Phi \underline{z}(0) + \Delta \underline{u}(0)] + C_N \}. \tag{3.21}$$

Again completing the square on $\underline{u}(0)$ gives

$$I_{N+1}[\underline{z}(0)] = \min_{\underline{\alpha}^- \le \underline{u}(0) \le \underline{\alpha}^+} \{ [\underline{u}(0) - \underline{u}'(0)]^T [\Delta^T (Q + P_N)\Delta + \Gamma][\underline{u}(0) - \underline{u}'(0)] + \underline{z}^T(0)\Phi^T(Q + P_N)\Phi \underline{z}(0) + \underline{z}^T(0)\Phi^T R_N + R_N^T \Phi \underline{z}(0) + C_N - [A_{N+1}\underline{z}(0) + B_{N+1}]^T [\Delta^T (Q + P_N)\Delta + \Gamma][A_{N+1}\underline{z}(0) + B_{N+1}] \}, \tag{3.22}$$

where

$$A_{N+1} = -[\Delta^T (Q + P_N)\Delta + \Gamma]^{-1} \Delta^T (Q + P_N)\Phi \tag{3.23}$$

as before, and

$$B_{N+1} = -[\Delta^T (Q + P_N)\Delta + \Gamma]^{-1} \Delta^T R_N \tag{3.24}$$

$$\underline{u}'(0) = A_{N+1}\underline{z}(0) + B_{N+1} . \tag{3.25}$$

The next step is to choose the $\underline{u}(0)$ that minimizes Eq. (3.22). This is easy when $u(0)$ is a scalar or when the distribution matrix Δ is an (m x m) diagonal matrix--an unlikely possibility. For the rest of this chapter and in Chapter IV, $u(0)$ will be considered a scalar, that is, there is only one controlling input to the system.

Since $[\Delta^T(Q + P_N)\Delta + \Gamma]$ is supposedly nonsingular (it is in fact a positive scalar), the minimum of Eq. (3.22) occurs at

$$u(0) = \begin{cases} \alpha^+ & \text{if } u'(0) > \alpha^+ \\ u'(0) & \text{if } \alpha^- \leq u'(0) \leq \alpha^+ \\ \alpha^- & \text{if } u'(0) < \alpha^- \end{cases} \tag{3.26}$$

The final step is to derive recurrence relations for P_{N+1}, R_{N+1}, and C_{N+1}. The existence of these relations gives the necessary proof that the form assumed for $I_N[\underline{z}(0)]$ is correct. Proof that the form for $I_1[\underline{z}(0)]$ is correct is the same as in Sec. A and will not be repeated.

The recurrence relations are different depending on whether or not $u(0)$ is saturated. When $u(0)$ is unsaturated, that is, when $\alpha^- \leq u'(0) \leq \alpha^+$, the relations can be obtained by equating (3.20) with (3.22) along with $u(0) = u'(0)$.

$$\begin{aligned} I_{N+1}[\underline{z}(0)] &= \underline{z}^T(0)\Phi^T(Q + P_N)\Phi\underline{z}(0) + \underline{z}^T(0)\Phi^T R_N + R_N^T\underline{z}(0) + C_N \\ &\quad - [A_{N+1}\underline{z}(0) + B_{N+1}]^T[\Delta^T(Q + P_N)\Delta + \Gamma][A_{N+1}\underline{z}(0) + B_{N+1}] \\ &= \underline{z}^T(0)P_{N+1}\underline{z}(0) + \underline{z}^T(0)R_{N+1} + R_{N+1}^T\underline{z}(0) + C_{N+1}. \end{aligned} \tag{3.27}$$

Thus the recurrence relations when $u(0)$ is unsaturated are

$$P_{N+1} = \Phi^T(Q + P_N)(\Phi + \Delta A_{N+1}) \tag{3.28}$$

$$R_{N+1} = (\Phi + \Delta A_{N+1})^T R_N \tag{3.29}$$

$$C_{N+1} = C_N + R_N^T \Delta B_{N+1}, \tag{3.30}$$

where the simplifying relations

$$-A_{N+1}^T[\Delta^T(Q + P_N)\Delta + \Gamma] = \Phi^T(Q + P_N)\Delta \tag{3.31}$$

and

$$-B_{N+1}^T[\Delta^T(Q + P_N)\Delta + \Gamma] = R_N^T\Delta \tag{3.32}$$

have been used.

When $u(n)$ is unbounded, these recurrence relations reduce, as they must, to that given in the unbounded case, Eq. (3.14). The equations for A_{N+1} and for P_{N+1} are the same as in the unbounded case. Since $R_0 = 0$, all $R_N = 0$. Since all $R_N = 0$, all $B_{N+1} = 0$. Finally, since $C_0 = 0$, all $C_N = 0$.

The recurrence relations when $u(0)$ is saturated can be determined by substituting $u(0) = \alpha$, where α represents either α^+ or α^-, into either Eq. (3.21) or (3.22) and equating the result with (3.20). Equating (3.21) with (3.20) gives

$$\begin{aligned} I_{N+1}[\underline{z}(0)] &= [\Phi\underline{z}(0) + \Delta\alpha]^T(Q + P_N)[\Phi\underline{z}(0) + \Delta\alpha] + \alpha^2\Gamma \\ &\quad + [\Phi\underline{z}(0) + \Delta\alpha]^T R_N + R_N^T[\Phi\underline{z}(0) + \Delta\alpha] + C_N \\ &= \underline{z}^T(0)P_{N+1}\underline{z}(0) + \underline{z}^T(0)R_{N+1} + R_{N+1}^T\underline{z}(0) + C_{N+1}. \end{aligned} \tag{3.33}$$

The recurrence relations when control is saturated are thus

$$P_{N+1} = \Phi^T(Q + P_N)\Phi \tag{3.34}$$

$$R_{N+1} = \Phi^T[R_N + (Q + P_N)\Delta\alpha] \tag{3.35}$$

$$C_{N+1} = C_N + \alpha^2[\Delta^T(Q + P_N)\Delta + \Gamma] + 2\alpha\Delta^T R_N. \tag{3.36}$$

By the same arguments used in the unsaturated control case, the form of $I_N[\underline{z}(0)]$ is shown to be correct by mathematical induction.

The principal equations derived in this section are summarized at the end of this chapter.

C. DISCUSSION

Beginning with a zero-stage process and calculating backward in time, as long as all the stages have optimally unsaturated control, the R_N, C_N, and B_{N+1} remain zero. The first stage backward in time that is saturated causes R_N and C_N to be nonzero, and they will remain nonzero for the rest of the stages.

If it were known a priori which terms of the optimal control sequence $u(0), u(1), \ldots, u(N-1)$ were saturated and which were not, the solution would proceed simply as in the unbounded case. Beginning with a one-stage process, A_1 and B_1 could be calculated. This requires no knowledge of whether or not the optimal control is saturated. Next, knowing whether the optimal control $u(0)$ equals α^+, α^-, or is unsaturated, P_1, R_1, and C_1 could be calculated. This computational scheme could be continued for as many stages as desired.

Unfortunately, nothing is known about the control sequence beforehand; thus the above computational scheme cannot be used. At each stage it is not known whether to use the recurrence relations for unsaturated or for saturated control. A computational method that does not require this a priori information is needed. Such a method will be discussed in Chapter IV.

The method described in the second paragraph of this section is still useful, however, and it has the advantage that it is exact. It can be used to perfect estimates of the optimal control obtained by other methods. For example, suppose the optimal control sequence was determined by a method requiring a discrete state space such as dynamic programming or the method described in the next chapter. Errors due to quantizing the state space will build up, and thus the true minimum and the true optimal unsaturated control will only be approximated. Now, however, it is known whether the control at each stage is saturated or not, and the simple computational scheme above can be applied to obtain the exact optimal control. Boundaries of all control regions of the examples in Chapter VI were checked in this manner.

SUMMARY OF PRINCIPAL EQUATIONS FOR SINGLE-INPUT CONTROL

Optimal return function

$$I_N[\underline{z}(0)] = \underline{z}^T(0)P_N\underline{z}(0) + \underline{z}^T(0)R_N + R_N^T\underline{z}(0) + C_N \tag{3.20}$$

Optimal control

$$u(0) = \begin{cases} \alpha^+ & \text{if } u'(0) > \alpha^+ \quad \text{(saturated)} \\ u'(0) & \text{if } \alpha^- \leq u'(0) \leq \alpha^+ \quad \text{(unsaturated)} \\ \alpha^- & \text{if } u'(0) < \alpha^- \quad \text{(saturated)} \end{cases} \tag{3.26}$$

Definitions

$$u'(0) = A_{N+1}\underline{z}(0) + B_{N+1} \tag{3.25}$$

$$A_{N+1} = -[\Delta^T(Q + P_N)\Delta + \Gamma]^{-1}\Delta^T(Q + P_N)\Phi \tag{3.23}$$

$$B_{N+1} = -[\Delta^T(Q + P_N)\Delta + \Gamma]^{-1}\Delta^T R_N \tag{3.24}$$

Recurrence relations:

Unsaturated control

$$P_{N+1} = \Phi^T(Q + P_N)(\Phi + \Delta A_{N+1}) \tag{3.28}$$

$$R_{N+1} = (\Phi + \Delta A_{N+1})^T R_N \tag{3.29}$$

$$C_{N+1} = C_N + R_N^T \Delta B_{N+1} \tag{3.30}$$

Saturated control

$$P_{N+1} = \Phi^T(Q + P_N)\Phi \tag{3.34}$$

$$R_{N+1} = \Phi^T[R_N + (Q + P_N)\Delta\alpha] \tag{3.35}$$

$$C_{N+1} = C_N + \alpha^2[\Delta^T(Q + P_N)\Delta + \Gamma] + 2\alpha\Delta^T R_N \tag{3.36}$$

Starting conditions

$$P_0 = 0, \qquad R_0 = 0, \qquad C_0 = 0.$$

IV. COMPUTATIONAL ASPECTS

This chapter describes a method of using the equations derived in Chapter III to determine the optimal control of any system described by Eqs. (2.1) and (2.2). This method, which requires a digital computer, calculates the optimal control from any point within a bounded region of state space for the infinite stage regulator problem. Useful facts pertaining to the actual computations are discussed.

Before describing the computing method recommended in this report, the straight dynamic programming approach will be briefly discussed for comparison.

A. DYNAMIC PROGRAMMING APPROACH

The basic dynamic programming approach to the problem is straightforward but requires a very large and very fast digital computer to solve for the optimal control of even small systems. This method repeatedly uses the fundamental functional equation of dynamic programming [Ref. 9] which, put into the form required for this problem, is

$$I_{N+1}[\underline{z}(0)] = \min_{\underline{\alpha}^- \leq \underline{u}(0) \leq \underline{\alpha}^+} \{J_1[\underline{z}(1)] + I_N[z(1)]\}, \tag{4.1}$$

where

$$J_1[\underline{z}(1)] = \underline{z}^T(1)Q\underline{z}(1) + \underline{u}^T(0)\Gamma\underline{u}(0), \tag{4.2}$$

and $I_N[\underline{z}(1)]$ is the minimum cost associated with initial condition $\underline{z}(1)$. Equation (4.1) is recognized as being the same as Eq. (3.19).

Although only the single-input case is being considered in this chapter, functions of the control $\underline{u}(n)$ will be written in vector-matrix form for use later in this report and for future work. Of course in the single-input case the last term of Eq. (4.2) is simply $u^2(0)\Gamma$.

In words, Eq. (4.1) states that the minimum cost $I_{N+1}[\underline{z}(0)]$ from initial state $\underline{z}(0)$ is the minimum over the allowable values of the control $\underline{u}(0)$ of the sum of the cost of the first step, which takes the state to $\underline{z}(1)$, plus the minimum cost of being in state $\underline{z}(1)$.

Before computing, both the state space and the control space must be quantized; that is, a discrete set of values is chosen over which the calculations are to be made. This set must be dense enough to prevent errors from accumulating during the calculations as the result of interpolation.

The calculation is divided into two parts: First the $I_N[\underline{z}(0)]$, called the optimal return functions, are calculated backward in time for all N and all $\underline{z}(0)$. Second, if actual optimal trajectories are desired, these are calculated forward in time using the optimal control calculated in the first part.

The first part of the calculation is the time-consuming part. Beginning with $I_0[\underline{z}(1)] = 0$ for all $\underline{z}(1)$, $I_1[\underline{z}(0)]$ is calculated from Eqs. (4.1) and (4.2) and the state-transition equation

$$\underline{z}(1) = \Phi\underline{z}(0) + \Delta\underline{u}(0). \tag{4.3}$$

For a given value of $\underline{z}(0)$ and for each value of $\underline{u}(0)$, $J_1[\underline{z}(1)]$ is calculated and the minimum is stored as $I_1[\underline{z}(1)]$. The optimal value of $\underline{u}(0)$ is also stored. This calculation is performed for each $\underline{z}(0)$.

Now that the values of $I_1[\underline{z}(1)]$ are known for all $\underline{z}(1)$, the $I_2[\underline{z}(0)]$ can be calculated, again using Eqs. (4.1) and (4.2) along with the state-transition equation (4.3). Since the $\underline{z}(1)$ calculated from $\underline{z}(0)$ by the state-transition equation will probably not be one of the discrete values for which the $I_1[\underline{z}(1)]$ were calculated, the correct value of $I_1[\underline{z}(1)]$ to use in Eq. (4.1) must be found by interpolation. It is the interpolation that causes the most significant errors to arise in the computation. Higher order than linear interpolation can be used, but since the interpolation must be done a very great number of times the computing time is increased significantly.

The process described in the last paragraph is continued until the optimal return functions and the optimal control for the desired N stages are calculated. In the case of the infinite stage regulator problem, stages must be calculated until the optimal control for each $\underline{z}(0)$ at stage $N+1$ is the same as the optimal control for each $\underline{z}(0)$ at stage N. This may require very many stages of calculation.

The fast memory requirements at each stage are three words for each value of $\underline{z}(0)$: $I_{N+1}[\underline{z}(0)]$, $I_N[\underline{z}(0)]$, and the optimal $\underline{u}(0)$ for stage N+1. Thus, for example, a two-dimensional problem with 100 values of z_1 and 100 values of z_2 would require 30,000 words of fast memory storage. This is approaching the limit of present-day computers. A three-dimensional problem with 100 points to each dimension would require 3,000,000 words of storage, thus exceeding the limit of present computers--a difficulty often referred to as the "curse of dimensionality."

The method discussed next for computing the special problem considered in this report requires far less computing storage and computing time than does straight dynamic programming.

B. A COMPUTING METHOD

The basic dynamic programming algorism makes no use of the recurrence relations derived in Chapter III. By taking advantage of this additional knowledge about the solution, considerable savings can be made in both computer time and memory, making it possible to solve much larger problems.

To facilitate the discussion of the computing method, which involves calculating regions of optimal control, several definitions will be made:

1. Region of linear control. In the infinite stage regulator problem there exists a region about the origin in state space where the control for the first and all future stages is unsaturated. Such a region will always exist if the plant is controllable, since in the unbounded control case the control is a linear function of the states and is zero at the origin. This region will be called the region of linear control, or simply the linear region.

2. Region of first saturation. If $\underline{z}(0)$ is not in the region of linear control, at least one stage before the state-space trajectory reaches the linear region will have saturated control. The first stage backward in time (or the last stage forward in time) that is saturated will be called the region of first saturation.

3. Unsaturated region. Any region where the control $u(0)$ is given by the equation $u(0) = u'(0) = A_{N+1}\underline{z}(0) + B_{N+1}$ will be called an unsaturated region. The region of linear control is an unsaturated region, but there will be others. Although the control in any unsaturated region is linear, the term "linear region" will refer only to the region of linear control.

4. Saturated region. This is a region where the control is either $u(0) = \alpha^+$ or $u(0) = \alpha^-$. Throughout, the term α will be used to denote either α^+ or α^-. Saturated regions will also be referred to as alpha-plus regions or alpha-minus regions.

5. $\underline{z}(1)$ region. This refers to a region that has already been calculated, and from which new regions will be calculated.

6. $\underline{z}(0)$ region. A $\underline{z}(0)$ region is one which is being presently calculated from a $\underline{z}(1)$ region. Regions are calculated backward in time as in dynamic programming; thus a $\underline{z}(0)$ region is calculated from a $\underline{z}(1)$ region. (The actual trajectories are, of course, from a $\underline{z}(0)$ region to the $\underline{z}(1)$ region from which it was calculated.)

The method to be described in detail is basically as follows: First the optimal feedback coefficients A_{N+1} for the infinite stage regulator problem with unbounded control are calculated. Once A_{N+1} is known, the region of linear control can be computed. Using the same A_{N+1} the two regions of first saturation are calculated. From each of these regions of first saturation are calculated an alpha-plus region, an unsaturated region, and an alpha-minus region. Further regions are calculated from each of these last regions, and the process is continued until all the state space of interest is covered with regions.

In essence, assuming N stages are being calculated backward from the linear region, this method considers all possible control sequences $u(0), u(1), \ldots, u(N-1)$, and determines the optimal sequence for each point in state space. Since at each stage the control can take one of three values--α^+, $u'(0)$, or α^---it might be thought that this method requires considering 3^N possible control sequences, a staggering possibility. In practice, the number of control sequences considered is far less. Most of the sequences will be found to be optimal for no points in state space, and these sequences can be dropped from further consideration as soon as they are discovered. The method described here determines these nonoptimal sequences at the earliest possible time during the computing.

As in dynamic programming, the state space must be quantized. However, the control is determined by the formulas of Chapter III and is not quantized.

A flow diagram of the computing is given in Fig. 3. First the optimal control feedback coefficients A_{N+1} for the linear region are calculated along with the corresponding P_N matrix, using the recurrence relations for unsaturated control. Beginning with $P_0 = 0$, A_1 is calculated. From A_1, P_1 is calculated. From P_1, A_2 is calculated. This iterative procedure is continued until the A_{N+1} converge to a limit. That these A_{N+1} will converge is discussed by Kalman [Ref. 11] and Gunckel [Ref. 12]. The unsaturated control recurrence relations for R_{N+1}, C_{N+1}, and the equation for B_{N+1}, show these to be zero for all N, since $R_0 = 0$ and $C_0 = 0$.

Second, the region of linear control is calculated. The optimal control formulas for this region are

$$u(n) = u'(n) = A_{N+1}\underline{z}(n) \tag{4.4}$$

$$\alpha^- \leq u'(n) \leq \alpha^+ \tag{4.5}$$

for all n, where the A_{N+1} is that calculated in the first step. Two bounds on this region can be found immediately by setting $u(n)$ in Eq. (4.4) equal to α^+ and α^-. Thus two bounds are

$$\alpha^+ = A_{N+1}\underline{z}(n)$$

$$\alpha^- = A_{N+1}\underline{z}(n). \tag{4.6}$$

For each $\underline{z}(0)$ on and within the boundaries (4.6) calculate

$$\underline{z}(1) = (\Phi + \Delta A_{N+1})\underline{z}(0). \tag{4.7}$$

Only those $\underline{z}(0)$ which determine $\underline{z}(1)$ that are on and within the boundaries (4.6) can be in the region of linear control. From each $\underline{z}(0)$ within the boundaries (4.6), enough points forward in time must be calculated to ensure that the $\underline{z}(0)$ is actually in the linear region. In the two-dimensional examples of Chapter VI, where $\alpha^+ = -\alpha^-$, only $\underline{z}(0)$ and $\underline{z}(1)$ both needed to be within the boundaries (4.6). In general more stages must be calculated.

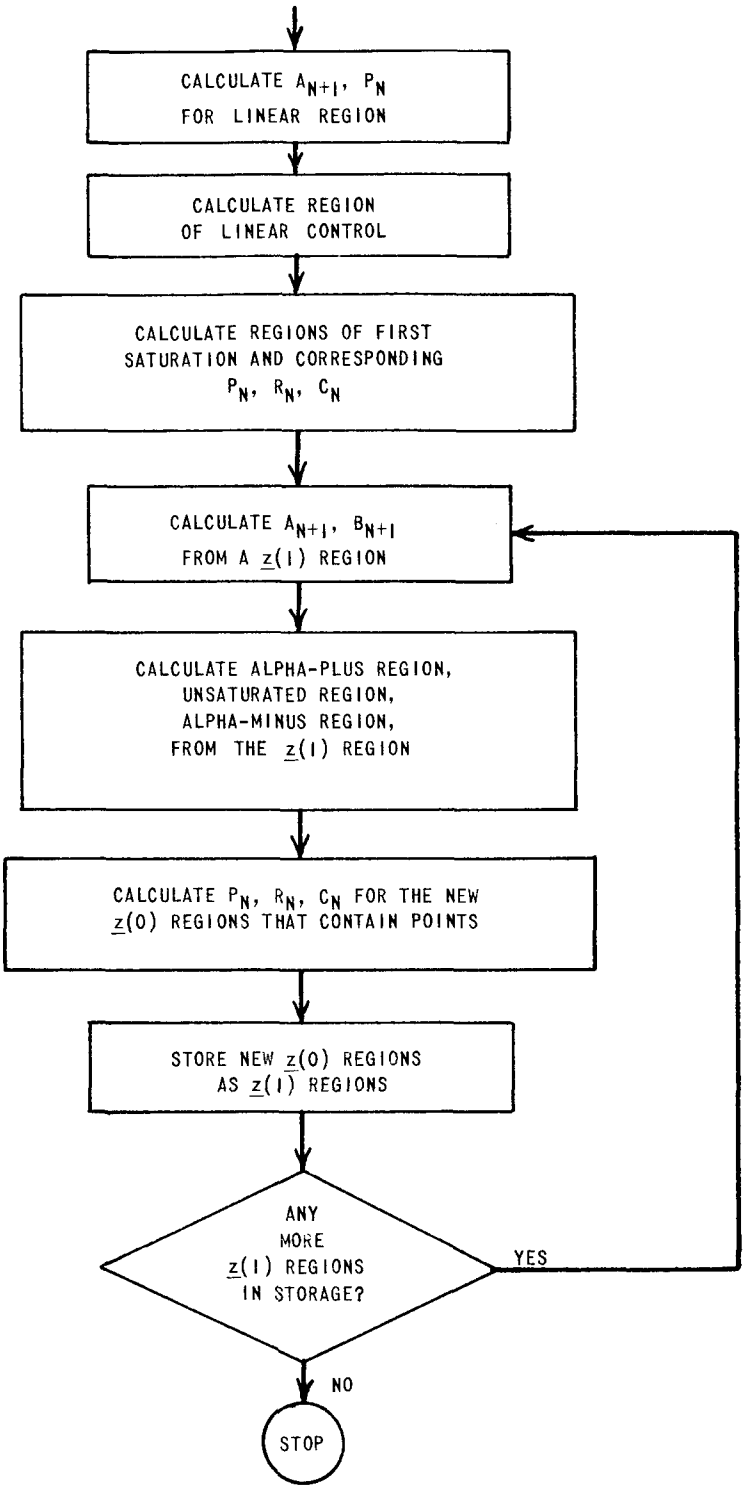

FIG. 3. FLOW DIAGRAM OF COMPUTING METHOD.

The third step is to calculate the two regions of first saturation. These are the $\underline{z}(0)$ regions that go optimally into the linear region with $u = \alpha$. Thus the alpha-plus region is defined by inequality (4.8) and Eq. (4.9):

$$u'(0) = A_{N+1}\underline{z}(0) > \alpha^+ \tag{4.8}$$

$$\underline{z}(0) = \Phi^{-1}[\underline{z}(1) - \Delta\alpha^+], \tag{4.9}$$

where $\underline{z}(1)$ is in the linear region. The A_{N+1} is the same as that used in calculating the linear region, since it is derived from the same P_N. The alpha-minus region is defined in a similar manner. If $\alpha^+ = -\alpha^-$, the alpha-minus region (and all regions derived from it) need not be calculated, since it is symmetric with respect to the origin to the alpha-plus region (and those derived from it). Finally, the P_N, R_N, and C_N are calculated for the regions of first saturation, using the saturated control recurrence relations.

The above steps are essentially initializing; the principal calculations now begin. There are now two $\underline{z}(1)$ regions from which to calculate--the two regions of first saturation. Consider the calculations from one of these. First the A_{N+1} and B_{N+1} are calculated using the P_N and R_N from the $\underline{z}(1)$ region. The optimal control for three $\underline{z}(0)$ regions--an alpha-plus region, an unsaturated region, and an alpha-minus region--is determined from these A_{N+1} and B_{N+1}. Each of these three regions must satisfy two relations as follows:

1. Alpha-plus region.

$$\underline{z}(0) = \Phi^{-1}[\underline{z}(1) - \Delta\alpha^+] \tag{4.10}$$

$$u'(0) = A_{N+1}\underline{z}(0) + B_{N+1} > \alpha^+ \tag{4.11}$$

2. Unsaturated region.

$$\underline{z}(0) = (\Phi + \Delta A_{N+1})^{-1}[\underline{z}(1) - \Delta B_{N+1}] \tag{4.12}$$

$$\alpha^- \leq u'(0) = A_{N+1}\underline{z}(0) + B_{N+1} \leq \alpha^+ \qquad (4.13)$$

3. Alpha-minus region.

$$\underline{z}(0) = \Phi^{-1}[\underline{z}(1) - \Delta\alpha^-] \qquad (4.14)$$

$$u'(0) = A_{N+1}\underline{z}(0) + B_{N+1} < \alpha^- \qquad (4.15)$$

In each of these equations $\underline{z}(1)$ is in the $\underline{z}(1)$ region, and for a $\underline{z}(0)$ to be in the new $\underline{z}(0)$ region, both the equation and the inequality for that region must be satisfied. Most of the regions calculated will be found to contain no states $\underline{z}(0)$. It is for this reason that there are considerably less than 3^N regions to consider.

The Eqs. (4.10), (4.12), and (4.14) are written as though $\underline{z}(0)$ will always be calculated from $\underline{z}(1)$ through an inverse relation. It is of course equally possible to calculate $\underline{z}(1)$ from $\underline{z}(0)$ by

$$\underline{z}(1) = \Phi\underline{z}(0) + \Delta u(0) \qquad (4.16)$$

for all $\underline{z}(0)$ in the quantized state space and keep only those $\underline{z}(0)$ for which the corresponding $\underline{z}(1)$ is in the desired $\underline{z}(1)$ region.

There are advantages and disadvantages for both methods of computing. Calculating $\underline{z}(1)$ from $\underline{z}(0)$ is easier because no "holes" can develop in the $\underline{z}(0)$ region. (Holes are points that belong within a region but are not calculated as being in the region.) However, because a very large percentage of the states $\underline{z}(0)$ will not be in the $\underline{z}(0)$ region, considerable computing time is consumed by computing $\underline{z}(1)$ from $\underline{z}(0)$.

Computing $\underline{z}(0)$ from $\underline{z}(1)$ consumes less computing time because less points are considered. Only those $\underline{z}(0)$ calculated from the $\underline{z}(1)$ in a particular $\underline{z}(1)$ region are considered. However, if the $\underline{z}(0)$ region contains more points than the $\underline{z}(1)$ region, holes will develop, and care must be taken to eliminate them. This is particularly a problem when calculating unsaturated regions. Also the points $\underline{z}(0)$ calculated from $\underline{z}(1)$ will be in a somewhat random order in the computer

memory and time must be taken to put them in some orderly and useful sequence.

Equations (4.10), (4.12), and (4.14) assume that the inverses of certain matrices exist. If the state variables have been chosen so that the minimum number necessary to completely characterize the system is used, the matrix Φ will be nonsingular; thus its inverse will exist. The other matrix assumed to be nonsingular is $(\Phi + \Delta A_{N+1})$. This matrix is nonsingular if Γ is not zero. However, if $\Gamma = 0$ this matrix will always be singular. The following proof will show an even stronger result: If $\Gamma = 0$ and $(\Phi + \Delta A_{N+1})$ has dimension (m x m), and Δ has rank q, (m is the dimension of $\underline{z}$ and q is ordinarily the dimension of $\underline{u}$,) then $(\Phi + \Delta A_{N+1})$ has rank no greater than m - q.

The proof is as follows. Consider a square matrix M of dimension m. If a nontrivial vector $\underline{c}$ can be found such that $\underline{c}^T M = \underline{0}$, then by definition M is singular. If there exist q nontrivial linearly independent vectors $\underline{c}$ such that there are q linearly independent vector equations $\underline{c}^T M = \underline{0}$, then q of the columns of M are linear combinations of the other m - q columns. The kernal of M is at least q and its rank is no greater than m - q.

The q nontrivial linearly independent vectors that show the matrix $(\Phi + \Delta A_{N+1})$ has rank no greater than m - q are the columns of $(Q + P_N)\Delta$. Thus

$$\Delta^T(Q + P_N)(\Phi + \Delta A_{N+1}) = \Delta^T(Q + P_N)\{\Phi - \Delta[\Delta^T(Q + P_N)\Delta]^{-1}\Delta^T(Q + P_N)\Phi\}$$

$$\equiv 0. \qquad (4.17)$$

This singularity can therefore be predicted in advance and the computer program written accordingly.

The next step is to calculate the P_N, R_N, and C_N for the $\underline{z}(0)$ regions just calculated that actually contain points. Regions that are found to contain no points are ignored entirely. The new $\underline{z}(0)$ regions are now stored in the fast memory as new $\underline{z}(1)$ regions. The old $\underline{z}(1)$ region can now be discarded.

The output can include a description of the points in the region, the type of region (alpha-plus, unsaturated, or alpha-minus), and the A_{N+1}, B_{N+1}, P_N, R_N, and C_N. The optimal cost can also be calculated and written.

This process of calculating $\underline{z}(0)$ regions is continued until the entire state space of interest is covered with regions.

C. DISCUSSION

This new program runs much more quickly than straight dynamic programming because the optimal control for each point is known from the recurrence relations. The memory requirements are also much smaller since only the $\underline{z}(0)$ regions and a single $\underline{z}(1)$ region need be in the fast memory at one time. It is convenient, however, to store all unused $\underline{z}(1)$ regions in the fast memory. Because of greater speed and less storage requirements, this new program can handle problems of larger dimension than can be run with straight dynamic programming. There is still, however, a limit to the size problem that can be run. A comparison of memory requirements is given for a specific example in Chapter VI. The restriction that the control be a scalar is removed in the next chapter.

The question of whether regions computed in this manner will overlap is still open. Such an overlap did not occur in any of the examples of Chapter VI. If, after computing, some regions do overlap, a comparison of the optimal costs from these regions can be made using their respective P_N's, R_N's, and C_N's, thus enabling the true optimal control to be chosen.

V. TWO-INPUT CONTROL

So far, only the solution to the single-input case has been completed in detail. This chapter extends these results to the case in which the control $\underline{u}(n)$ is a two-dimensional vector. The solution is considerably complicated by the fact that one of the inputs may be saturated while the other is not. Although they are not discussed here, this chapter indicates the extensions and changes that must be made when the control has dimension higher than two.

A. THE PROBLEM

The description of the system and performance criterion is the same as given in Chapter II. The control $\underline{u}(n)$ and its bounds $\underline{\alpha}^+$ and $\underline{\alpha}^-$ are now two-dimensional vectors.

$$\underline{u}(n) = \begin{bmatrix} u_1(n) \\ u_2(n) \end{bmatrix}, \qquad \underline{\alpha}^+ = \begin{bmatrix} \alpha_1^+ \\ \alpha_2^+ \end{bmatrix}, \qquad \underline{\alpha}^- = \begin{bmatrix} \alpha_1^- \\ \alpha_2^- \end{bmatrix}. \tag{5.1}$$

Control is limited by the vector inequality

$$\underline{\alpha}^- \leq \underline{u}(n) \leq \underline{\alpha}^+. \tag{5.2}$$

The problem is: Given any initial condition $\underline{z}(0)$, find the optimal control-vector sequence $\underline{u}(0)$, $\underline{u}(1)$, $\underline{u}(2)$, ... that minimizes the performance index $J_\infty[\underline{z}(0)]$.

B. THE SOLUTION

The equations for $I_{N+1}[\underline{z}(0)]$ derived in Chapter III up through Eq. (3.25) were written in vector notation so that they could be used in this chapter. Equation (3.21) is written here as the starting point of the solution:

$$I_{N+1}[\underline{z}(0)] = \min_{\underline{\alpha}^- \leq \underline{u}(0) \leq \underline{\alpha}^+} \{ [\Phi\underline{z}(0) + \Delta\underline{u}(0)]^T (Q + P_N) [\Phi\underline{z}(0) + \Delta\underline{u}(0)] + \underline{u}^T(0) \Gamma \underline{u}(0) + [\Phi\underline{z}(0) + \Delta\underline{u}(0)]^T R_N + R_N^T [\Phi\underline{z}(0) + \Delta\underline{u}(0)] + C_N \}. \tag{5.3}$$

As in Chapter III the square is completed on the control vector $\underline{u}(0)$. The result is Eq. (3.22):

$$I_{N+1}[\underline{z}(0)] = \min_{\underline{\alpha}^- \leq \underline{u}(0) \leq \underline{\alpha}^+} \{[\underline{u}(0) - \underline{u}'(0)]^T[\Delta^T(Q + P_N)\Delta + \Gamma][\underline{u}(0) - \underline{u}'(0)]$$
$$+ \underline{z}^T(0)\Phi^T(Q + P_N)\Phi\underline{z}(0) + \underline{z}^T(0)\Phi^T R_N + R_N^T\Phi\underline{z}(0) + C_N$$
$$- [A_{N+1}\underline{z}(0) + B_{N+1}]^T[\Delta^T(Q + P_N)\Delta + \Gamma][A_{N+1}\underline{z}(0) + B_{N+1}]\}, \tag{5.4}$$

where A_{N+1}, B_{N+1}, and $\underline{u}'(0)$ are given by Eqs. (3.23) through (3.25).

The minimum without regard to bounds occurs at $\underline{u}(0) = \underline{u}'(0)$. If the resulting $\underline{u}(0)$ satisfies the vector inequality (5.2), then $\underline{u}(0) = \underline{u}'(0)$ is the optimal control. However, if one or both of the elements of $\underline{u}'(0)$ are out of bounds, the situation is much complicated.

Before presenting a careful algebraic discussion of a method for finding the minimum of Eq. (5.4), a more intuitive geometrical discussion will be given.

C. GEOMETRICAL DISCUSSION OF THE MINIMUM

Figures 4a through 4f show the two-dimensional control space. Each point represents a particular control (u_1, u_2). The rectangle, whose sides are given by $u_1 = \alpha_1^+$, $u_1 = \alpha_1^-$, $u_2 = \alpha_2^+$, and $u_2 = \alpha_2^-$, bounds the region of allowable control.

Geometrically, a positive definite quadratic function in two variables is an ellipse. In each figure are drawn concentric ellipses, which are loci of constant $J_{N+1}[\underline{z}(0)]$. The value of $J_{N+1}[\underline{z}(0)]$ decreases as the ellipse size decreases. The absolute minimum occurs at the center of the ellipses, which has coordinates (u_1', u_2').

The geometrical problem then is to find the point in the control space that is both on the smallest possible ellipse and in or on the rectangle. Algebraically this is the same problem as expressed by the now-familiar equation

$$I_{N+1}[\underline{z}(0)] = \min_{\underline{\alpha}^- \leq \underline{u}(0) \leq \underline{\alpha}^+} \{J_{N+1}[\underline{z}(0)]\}. \tag{5.5}$$

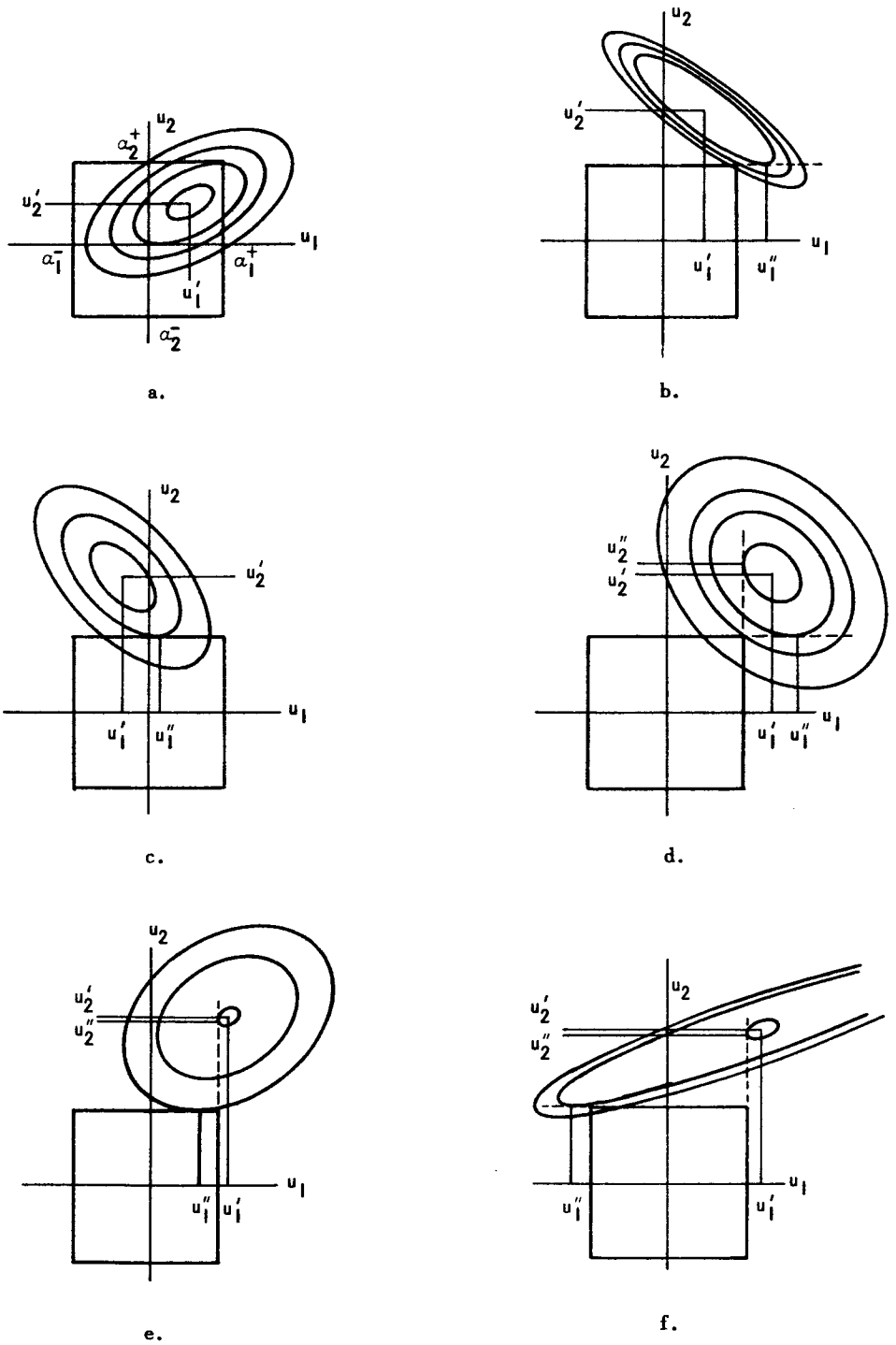

FIG. 4. GEOMETRICAL DESCRIPTION OF MINIMUM.

If the center of the ellipses is in the rectangle, as shown in Fig. 4a, then the minimum occurs at (u_1', u_2'). If the center is outside the rectangle, as shown in Figs. 4b through 4f, then it is clear that the minimum occurs on the boundary, since any control within the rectangle is on a larger ellipse than one either tangent to the boundary or touching a corner.

Figures 4b and 4c show the case where one u_i', in this case u_2', is greater than its bound, while the other u_i' is within its bounds. In both cases the optimal control of u_2 is α_2^+, but the optimal value of u_1 can be anything. To determine u_1, the optimal value of $u_2 = \alpha_2^+$ is substituted into Eq. (5.3) or (5.4), and by completing the square on u_1 the value of u_1 that minimizes the function is found. This value, called u_1'', may or may not be in bounds. The optimal u_1 is then

$$u_1 = \begin{cases} \alpha_1^+ & \text{if } u_1'' > \alpha_1^+ \quad \text{(Fig. 4b)} \\ u_1'' & \text{if } \alpha_1^- \leq u_1'' \leq \alpha_1^+ \quad \text{(Fig. 4c)} \\ \alpha_1^- & \text{if } u_1'' < \alpha_1^- \end{cases} \tag{5.6}$$

Note that the value of u_1' does not in any way indicate the optimal value of u_1.

Figures 4d through 4f show cases where both u_1' and u_2' are outside the bounds. In these cases all that can be said without further calculation is that at least one of the u_i' gives the optimal control. Geometrically this means that the optimal control is on one of the two boundaries nearest the center of the ellipse, a fact that will be proved algebraically in the next section. Since it is not known which u_i' gives the correct result, both u_1'' and u_2'' must be calculated. Assume u_1' and u_2' are greater than α_1^+ and α_2^+ respectively as in the figures. Then u_1'' is calculated as the optimal value of u_1 (neglecting saturation of u_1) with $u_2 = \alpha_2^+$, and u_2'' is likewise calculated. Since only one of the assumptions made in calculating the u_i'' was necessarily correct, only one of the u_i'' is necessarily correct. However, as shown in the next section, both u_i'' calculated determine correctly their respective optimal u_i.

If the correct optimal value of u_2 is α_2^+ as shown in Figs. 4d through 4f, then the value of u_1'' calculated is correct. The optimal u_1 is thus

$$u_1 = \begin{cases} \alpha_1^+ & \text{if } u_1'' > \alpha_1^+ \quad \text{(Fig. 4d)} \\ u_1'' & \text{if } \alpha_1^- \leq u_1'' \leq \alpha_1^+ \quad \text{(Fig. 4e)} \\ \alpha_1^- & \text{if } u_1'' < \alpha_1^- \quad \text{(Fig. 4f)} \end{cases} \tag{5.7}$$

Note that even though the value of u_1' suggests that the optimal value of u_1 is α_1^+, the real optimal value can be far different, even α_1^-.

The next section algebraically proves that the u_i'', where calculated, give the optimal values of u_i in all cases.

D. ALGEBRAIC DETERMINATION OF THE MINIMUM

To simplify the notation in this section, consider only the part of $J_{N+1}[\underline{z}(0)]$ that is quadratic in $\underline{u}$. This is

$$\xi = (\underline{u} - \underline{u}')^T K(\underline{u} - \underline{u}') \tag{5.8}$$

where

$$(\underline{u} - \underline{u}') = \begin{bmatrix} (u_1 - u_1') \\ (u_2 - u_2') \end{bmatrix}, \tag{5.9}$$

and

$$K = [\Delta^T(Q + P_N)\Delta + \Gamma] = \begin{bmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{bmatrix} \tag{5.10}$$

K is a positive definite symmetric matrix, and hence the smallest value ξ can have is zero, which occurs only at $\underline{u} = \underline{u}'$. If $\underline{\alpha}^- \leq \underline{u}' \leq \underline{\alpha}^+$, the optimum value of $\underline{u}$ is clearly $\underline{u} = \underline{u}'$.

The quantities u_i'' have a somewhat more general meaning in this section than in the last section. For example u_1'', is the optimum value

of u_1 (ignoring saturation) for any given value of u_2, not just α_2. It is derived by completing the square on ξ with respect to u_1. Thus

$$\xi = (u_1 - u_1')^2 k_{11} + 2(u_1 - u_1')(u_2 - u_2')k_{12} + (u_2 - u_2')^2 k_{22} \tag{5.11}$$

$$\xi = (u_1 - u_1'')^2 k_{11} - u_1''^2 k_{11} + \text{terms not involving } u_1. \tag{5.12}$$

By equating (5.12) with (5.11) an equation for u_1'' is determined:

$$u_1'' = u_1' - \frac{(u_2 - u_2')k_{12}}{k_{11}} . \tag{5.13}$$

Likewise,

$$u_2'' = u_2' - \frac{(u_1 - u_1')k_{12}}{k_{22}} . \tag{5.14}$$

The optimal control is determined when four equations are simultaneously satisfied. These are Eqs. (5.13) and (5.14) along with

$$u_1 = \begin{cases} \alpha_1^+ & \text{if } u_1'' > \alpha_1^+ \\ u_1'' & \text{if } \alpha_1^- \le u_1'' \le \alpha_1^+ \\ \alpha_1^- & \text{if } u_1'' < \alpha_1^- \end{cases} \tag{5.15}$$

$$u_2 = \begin{cases} \alpha_2^+ & \text{if } u_2'' > \alpha_2^+ \\ u_2'' & \text{if } \alpha_2^- \le u_2'' \le \alpha_2^+ \\ \alpha_2^- & \text{if } u_2'' < \alpha_2^- \end{cases} \tag{5.16}$$

These four equations can be solved on an analog computer very simply, but this is not much help here. The following proofs show a simple way to determine the optimal values of u_1 and u_2. In the following discussion u_i' will be called not admissible when either $u_i' > \alpha_i^+$ or $u_i' < \alpha_i^-$. Otherwise it will be called admissible.

Two cases need to be considered: (1) only one u_i' is not admissible, and (2) neither u_i' is admissible. Before beginning two inequality relations involving the elements of a symmetric positive definite matrix must be stated. These can be found in nearly any book on matrix theory [Ref. 17].

$$k_{11} > 0, \qquad (\text{and } k_{22} > 0), \tag{5.17}$$

$$k_{11}k_{22} - k_{12}^2 > 0. \tag{5.18}$$

1. Case 1

The first case occurs when one u_i' is admissible while the other is not. For definiteness let

$$\alpha_1^- \leq u_1' \leq \alpha_1^+ \tag{5.19}$$

$$u_2' > \alpha_2^+. \tag{5.20}$$

It will now be shown that the optimal value of u_2 is α_2^+ regardless of the optimal value of u_1. From (5.13)

$$u_1'' - u_1' = \frac{[-(u_2 - u_2')]k_{12}}{k_{11}} \tag{5.21}$$

The quantity in brackets is positive regardless of the choice of u_2 because of inequality (5.20). Thus $(u_1'' - u_1')$ has the same sign as k_{12}.

If k_{12} is negative or zero, then $u_1'' \leq u_1'$, and either $u_1 = u_1''$ or $u_1 = \alpha_1^-$. Thus

$$u_1'' \leq u_1 \leq u_1'. \tag{5.22}$$

Therefore

$$0 \geq u_1 - u_1' \geq u_1'' - u_1'. \tag{5.23}$$

When k_{12} is negative, multiplying inequality (5.23) through by k_{12}/k_{22} reverses the inequality signs. The result is of course trivial when k_{12} is zero.

$$0 \leq \frac{(u_1 - u_1')k_{12}}{k_{22}} \leq \frac{(u_1'' - u_1')k_{12}}{k_{22}} = -\frac{(u_2 - u_2')k_{12}^2}{k_{11}k_{22}} . \qquad (5.24)$$

The right side of (5.24) is determined by using Eq. (5.21).

If k_{12} is positive or zero, the same steps outlined in the last paragraph can be taken, and the result is again (5.24).

Equation (5.24) shows that the last term in Eq. (5.14) is positive or zero, regardless of the sign of k_{12}. Using (5.24), Eq. (5.14) can be written as the following inequality:

$$u_2'' \geq u_2' + \frac{(u_2 - u_2')k_{12}^2}{k_{11}k_{22}} \qquad (5.25)$$

$$u_2'' \geq u_2' \left(1 - \frac{k_{12}^2}{k_{11}k_{22}}\right) + u_2 \frac{k_{12}^2}{k_{11}k_{22}} . \qquad (5.26)$$

Since $k_{12}^2/k_{11}k_{22} < 1$ as shown by (5.18), the term in parentheses in (5.26) is positive. Substituting (5.20) into (5.26) makes the inequality even stronger.

$$u_2'' > \alpha_2^+ \left(1 - \frac{k_{12}^2}{k_{11}k_{22}}\right) + u_2 \frac{k_{12}^2}{k_{11}k_{22}} . \qquad (5.27)$$

The only u_2 that satisfies both (5.27) and (5.16) is $u_2 = \alpha_2^+$. Substituting $u_2 = u_2''$ into (5.27) leads to a contradiction.

The conclusion is that if u_i' is not admissible and u_j' $(j \neq i)$ is admissible, the optimal value of u_i equals the nearest bound to u_i', while u_j is determined from u_j''.

2. Case 2

The second case occurs when neither u_1' nor u_2' is admissible. For definiteness let

$$u_1' > \alpha_1^+ \tag{5.28}$$

$$u_2' > \alpha_2^+ \tag{5.29}$$

Equations (5.13) and (5.14) written in the form of (5.21) show that both $(u_1'' - u_1')$ and $(u_2'' - u_2')$ have the same sign as k_{12}.

If k_{12} is positive or zero then

$$u_1'' \geq u_1' > \alpha_1^+ \tag{5.30}$$

$$u_2'' \geq u_2' > \alpha_2^+ \tag{5.31}$$

Thus the optimal control is $u_1 = \alpha_1^+$, $u_2 = \alpha_2^+$.

If k_{12} is negative the situation is much more complicated. The values of u_1' and u_2' given by (5.28) and (5.29) determine that either $u_1 = \alpha_1^+$ or $u_2 = \alpha_2^+$ or both. This will be proved next.

The proof assumes that both the optimal $u_2 < \alpha_2^+$ and $u_1'' < \alpha_1^+$ occur simultaneously, and arrives at a contradiction. Since both conditions cannot occur simultaneously, at least one u_i must equal α_i^+.

If $u_1'' < \alpha_1^+$, then either $u_1 = u_1''$ or $u_1 = \alpha_1^-$. Thus

$$u_1' > u_1 \geq u_1''. \tag{5.32}$$

Combining (5.32) with (5.13) gives

$$0 > u_1 - u_1' \geq - \frac{(u_2 - u_2')k_{12}}{k_{11}} . \tag{5.33}$$

The direction of the inequality is changed when (5.33) is multiplied through by the negative quantity k_{12}/k_{22}.

$$0 < \frac{(u_1 - u_1')k_{12}}{k_{22}} \leq - \frac{(u_2 - u_2')k_{12}^2}{k_{11}k_{22}} . \tag{5.34}$$

Using (5.34), Eq. (5.14) becomes inequality (5.35):

$$u_2'' \geq u_2' + \frac{(u_2 - u_2')k_{12}^2}{k_{11}k_{22}} \tag{5.35}$$

$$u_2'' \geq u_2' \left(1 - \frac{k_{12}^2}{k_{11}k_{22}}\right) + u_2 \frac{k_{12}^2}{k_{11}k_{22}} . \tag{5.36}$$

Since $u_2' > u_2$ by (5.29), inequality (5.36) becomes

$$u_2'' > u_2 \left(1 - \frac{k_{12}^2}{k_{11}k_{22}}\right) + u_2 \frac{k_{12}^2}{k_{11}k_{22}} = u_2 \tag{5.37}$$

which contradicts the original assumption that $u_2 < \alpha_2^+$ (and thus $u_2'' \leq u_2$). Thus the proof is complete.

Since it is not known which $u_i = \alpha_i^+$, calculate u_1'' on the assumption that $u_2 = \alpha_2^+$, and calculate u_2'' on the assumption that $u_1 = \alpha_1^+$. If both $u_i'' \geq \alpha_i^+$ then both assumptions were correct and the optimal control is determined.

At least one assumption was correct, thus at least one u_i'' was computed correctly. Assume $u_1 = \alpha_1^+$ is correct but $u_2 = \alpha_2^+$ is incorrect; then u_2'' is correct but not u_1''. The last step is to prove that $u_1'' > \alpha_1^+$ even though it was computed using an incorrect assumption. Thus

$$u_1'' = u_1' - \frac{(\alpha_2^+ - u_2')k_{12}}{k_{11}} \tag{5.38}$$

and

$$u_2' > \alpha_2^+ > u_2''. \tag{5.39}$$

Again since k_{12} is negative,

$$0 < \frac{(\alpha_2^+ - u_2')k_{12}}{k_{11}} < \frac{(u_2'' - u_2')k_{12}}{k_{11}} . \tag{5.40}$$

Substituting (5.40) into (5.38) gives

$$u_1'' > u_1' - \frac{(u_2'' - u_2')k_{12}}{k_{11}} = u_1' + \frac{(\alpha_1^+ - u_1')k_{12}^2}{k_{11}k_{22}} > \alpha_1^+. \tag{5.41}$$

To sum up, the above proofs show a simple way to calculate the control that minimizes the quadratic function given by Eq. (5.11):

1. Case 1. If u_i' is not admissible but u_j' is admissible, then u_i equals the bound nearest to u_i', and u_j is determined from u_j''. The value of u_j'' is calculated using the known optimal value of u_i, namely α_i^+ or α_i^-.

2. Case 2. If neither u_1' nor u_2' is admissible, then u_1'' is computed using the bound nearest to u_2' for u_2, and u_2'' is likewise computed. The optimal control is then determined from Eqs. (5.15) and (5.16).

E. OPTIMAL CONTROL FORMULAS AND RECURRENCE RELATIONS

The recurrence relations when neither control is saturated are the same as the unsaturated control recurrence relations given in Chapter III. When both controls are saturated, the relations are the same as the saturated control recurrence relations in Chapter III, though written in vector notation. Thus the only new recurrence relations are for the case in which one control is saturated and the other is not.

Assume the optimal control is given by $u_1(0) = u_1''(0)$ and $u_2(0) = \alpha_2$. The derivation of $u_1''(0)$ and the recurrence relations begins by completing the square on $u_1(0)$, assuming $u_2(0) = \alpha_2$ (where as usual α_2 represents either α_2^+ or α_2^-). Note that completing the square on $u_1(0)$ is not at all the same as completing the square on the vector $\underline{u}(0)$.

The elements of the control vector $\underline{u}(0)$ are separated in Eq. (5.3) by partitioning the Δ matrix as follows:

$$\Delta = (\Delta_1 \vdots \Delta_2), \tag{5.42}$$

where the Δ_i $(i = 1,2)$ are $(m \times 1)$ column matrices. With this partitioning, Eq. (5.3) is written as

$$I_{N+1}[\underline{z}(0)] = \min_{\alpha_1^- \le u_1(0) \le \alpha_1^+} \Big\{ u_1^2(0)\gamma_{11} + 2u_1(0)\alpha_2\gamma_{12} + \alpha_2^2\gamma_{22}$$

$$+ [\Phi\underline{z}(0) + \Delta_1 u_1(0) + \Delta_2\alpha_2]^T(Q + P_N)[\Phi\underline{z}(0) + \Delta_1 u_1(0) + \Delta_2\alpha_2]$$

$$+ [\Phi\underline{z}(0) + \Delta_1 u_1(0) + \Delta_2\alpha_2]^T R_N$$

$$+ R_N^T[\Phi\underline{z}(0) + \Delta_1 u_1(0) + \Delta_2\alpha_2] + C_N \Big\} . \qquad (5.43)$$

Completing the square on $u_1(0)$ gives

$$I_{N+1}[\underline{z}(0)] = \min_{\alpha_1^- \le u_1(0) \le \alpha_1^+} \Big\{ [u_1(0) - u_1''(0)]^2 [\Delta_1^T(Q + P_N)\Delta_1 + \gamma_{11}]$$

$$- [A'_{1_{N+1}}\underline{z}(0) + B'_{1_{N+1}}]^T[\Delta_1^T(Q + P_N)\Delta_1 + \gamma_{11}][A'_{1_{N+1}}\underline{z}(0) + B'_{1_{N+1}}]$$

$$+ [\Phi\underline{z}(0) + \Delta_2\alpha_2]^T(Q + P_N)[\Phi\underline{z}(0) + \Delta_2\alpha_2] + \alpha_2^2\gamma_{22}$$

$$+ [\Phi\underline{z}(0) + \Delta_2\alpha_2]^T R_N + R_N^T[\Phi\underline{z}(0) + \Delta_2\alpha_2] + C_N \Big\}, \qquad (5.44)$$

where

$$u_1''(0) = A'_{1_{N+1}}\underline{z}(0) + B'_{1_{N+1}} \qquad (5.45)$$

$$A'_{1_{N+1}} = - \frac{\Delta_1^T(Q + P_N)\Phi}{\Delta_1^T(Q + P_N)\Delta_1 + \gamma_{11}} \qquad (5.46)$$

$$B'_{1_{N+1}} = - \frac{\Delta_1^T R_N + \Delta_1^T(Q + P_N)\Delta_2\alpha_2 + \gamma_{12}\alpha_2}{\Delta_1^T(Q + P_N)\Delta_1 + \gamma_{11}} \qquad (5.47)$$

Equations for u_2'' and its associated $A'_{2_{N+1}}$ and $B'_{2_{N+1}}$ are determined in the same manner.

The recurrence relations for the case where $u_1 = u_1''$ and $u_2 = \alpha_2$ are determined by equating the assumed form of $I_{N+1}[\underline{z}(0)]$ given by

Eq. (3.20) with Eq. (5.44). These recurrence relations are

$$P_{N+1} = \Phi^T(Q + P_N)(\Phi + \Delta_1 A'_{1_{N+1}}) \tag{5.48}$$

$$R_{N+1} = \Phi^T[R_N + (Q + P_N)(\Delta_1 B'_{1_{N+1}} + \Delta_2\alpha_2)] \tag{5.49}$$

$$C_{N+1} = C_N + [\Delta_1^T R_N + \Delta_1^T(Q + P_N)\Delta_2\alpha_2 + \gamma_{12}\alpha_2]B'_{1_{N+1}} + 2\alpha_2\Delta_2^T R_N + \alpha_2^2[\Delta_2^T(Q + P_N)\Delta_2 + \gamma_{22}] \tag{5.50}$$

As shown in Chapter III, the existence of these recurrence relations shows that the form assumed for $I_{N+1}[\underline{z}(0)]$ is correct.

The equations and recurrence relations for the two-input control case are summarized at the end of this chapter.

F. COMPUTING METHOD

Computing proceeds as in Chapter IV with only a few changes. The first step is to calculate the region of linear control. Next the eight regions of first saturation are calculated. These are the regions computed from the linear region that have one or both controls saturated. From each of these regions are calculated nine more regions, regions with the nine possible combinations of the controls. Regions are computed in this manner until all the state space of interest is covered.

Certainly the two-dimensional control case will take much more computing time and storage than the one-dimensional case. The method outlined here could conceivably be extended to higher-dimensional control, but the complexity increases rapidly.

SUMMARY OF PRINCIPAL EQUATIONS FOR TWO-INPUT CONTROL

Optimal return function

$$I_N[\underline{z}(0)] = \underline{z}^T(0)P_N\underline{z}(0) + \underline{z}^T(0)R_N + R_N^T\underline{z}(0) + C_N \qquad (3.20)$$

Optimal control

$$\underline{u}(0) = \underline{u}'(0) \quad \text{if} \quad \underline{\alpha}^- \leq \underline{u}'(0) \leq \underline{\alpha}^+ \qquad \text{(unsaturated)}$$

$$\underline{u}(0) = \underline{\alpha} \quad \text{if} \quad u_1''(0) > \alpha_1^+ \ \underline{\text{or}} \ u_1''(0) < \alpha_1^- \ \underline{\text{and}}$$

$$u_2''(0) > \alpha_2^+ \ \underline{\text{or}} \ u_2''(0) < \alpha_2^- \qquad \text{(saturated)}$$

$$\left.\begin{aligned} u_i(0) &= \alpha_i \\ u_j(0) &= u_j''(0) \end{aligned}\right\} \quad \text{if} \quad \begin{cases} u_i''(0) > \alpha_i^+ \ \underline{\text{or}} \ u_i''(0) < \alpha_i^- \ \underline{\text{and}} \\ \alpha_j^- \leq u_j''(0) \leq \alpha_j^+ \end{cases} \qquad \text{(mixed)}$$

Definitions

$$\underline{u}'(0) = A_{N+1}\underline{z}(0) + B_{N+1} \qquad (3.25)$$

$$A_{N+1} = -[\Delta^T(Q + P_N)\Delta + \Gamma]^{-1}\Delta^T(Q + P_N)\Phi \qquad (3.23)$$

$$B_{N+1} = -[\Delta^T(Q + P_N)\Delta + \Gamma]^{-1}\Delta^T R_N \qquad (3.24)$$

$$\Delta = (\Delta_1 \vdots \Delta_2) \qquad (5.42)$$

$$u_j''(0) = A'_{j_{N+1}}\underline{z}(0) + B'_{j_{N+1}}, \qquad j = 1, 2 \qquad (5.45)$$

$$A'_{j_{N+1}} = -\frac{\Delta_j^T(Q + P_N)\Phi}{\Delta_j^T(Q + P_N)\Delta_j + \gamma_{jj}} \qquad (5.46)$$

$$B'_{j_{N+1}} = -\frac{\Delta_j^T R_N + \Delta_j^T(Q + P_N)\Delta_i\alpha_i + \gamma_{12}\alpha_i}{\Delta_j^T(Q + P_N)\Delta_j + \gamma_{jj}} \qquad (5.47)$$

Recurrence relations:

Unsaturated control

$$P_{N+1} = \Phi^T(Q + P_N)(\Phi + \Delta A_{N+1}) \tag{3.28}$$

$$R_{N+1} = (\Phi + \Delta A_{N+1})^T R_N \tag{3.29}$$

$$C_{N+1} = C_N + R_N^T \Delta B_{N+1} \tag{3.30}$$

Saturated control

$$P_{N+1} = \Phi^T(Q + P_N)\Phi \tag{3.34}$$

$$R_{N+1} = \Phi^T[R_N + (Q + P_N)\Delta\underline{\alpha}] \tag{3.35}$$

$$C_{N+1} = C_N + \underline{\alpha}^T[\Delta^T(Q + P_N)\Delta + \Gamma]\underline{\alpha} + 2\underline{\alpha}^T\Delta^T R_N \tag{3.36}$$

Mixed control (u_i saturated; u_j unsaturated)

$$P_{N+1} = \Phi^T(Q + P_N)(\Phi + \Delta_j A'_{j_{N+1}}) \tag{5.48}$$

$$R_{N+1} = \Phi^T[R_N + (Q + P_N)(\Delta_j B'_{j_{N+1}} + \Delta_i \alpha_i)] \tag{5.49}$$

$$C_{N+1} = C_N + [\Delta_j^T R_N + \Delta_j^T(Q + P_N)\Delta_i\alpha_i + \gamma_{12}\alpha_i]B'_{j_{N+1}} + 2\alpha_i\Delta_i^T R_N$$

$$+ \alpha_i^2[\Delta_i^T(Q + P_N)\Delta_i + \gamma_{ii}] \tag{5.50}$$

Starting conditions

$$P_0 = 0, \qquad R_0 = 0, \qquad C_0 = 0.$$

VI. EXAMPLES

In this chapter four examples of optimal control systems computed using the method described in Chapter IV are presented. Some optimal trajectories in state space are also shown. In the final section the synthesis of the systems is discussed.

A. EXAMPLE A

For the first example consider the space vehicle described in Chapter II. The optimal control that minimizes the attitude sum-squared-error from any initial attitude error and its rate of change is to be found. Since power consumption is an important design consideration, the total energy used in controlling the vehicle is charged by including sum-squared-control in the performance criterion.

All parameters are normalized to unity, and the sampling interval is arbitrarily set at $\tau = 1$. The resulting system is shown in Fig. 5.

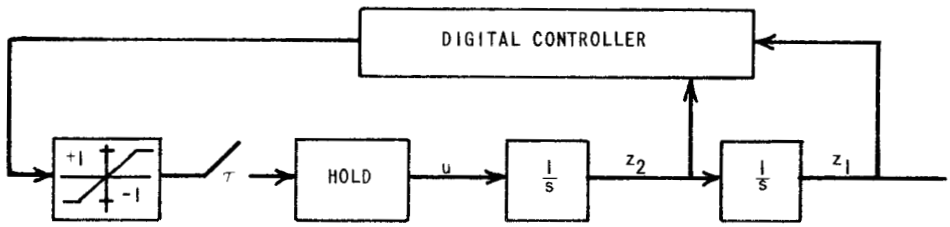

FIG. 5. BLOCK DIAGRAM OF SYSTEM FOR EXAMPLES A-C.

The state-transition equations, as derived in Chapter II, are

$$\underline{z}(n+1) = \begin{bmatrix} 1.0 & 1.0 \\ 0.0 & 1.0 \end{bmatrix} \underline{z}(n) + \begin{bmatrix} 0.5 \\ 1.0 \end{bmatrix} u(n), \qquad (6.1)$$

where control is limited by

$$-1.0 \leq u(n) \leq +1.0. \qquad (6.2)$$

The performance index is

$$J_{\infty}[\underline{z}(0)] = \sum_{n=1}^{\infty} [z_1^2(n) + u^2(n-1)] \tag{6.3}$$

and thus

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \Gamma = 1. \tag{6.4}$$

The optimal control for this example is shown in Fig. 6. This figure shows the state space divided into three main parts. In the upper area of the figure the optimal control is $u = -1$. This area is composed of all the alpha-minus regions that were calculated using the method described in Chapter IV. The boundaries of these regions are not shown in the figure, since they represent information that is unnecessary to the synthesis of the system.

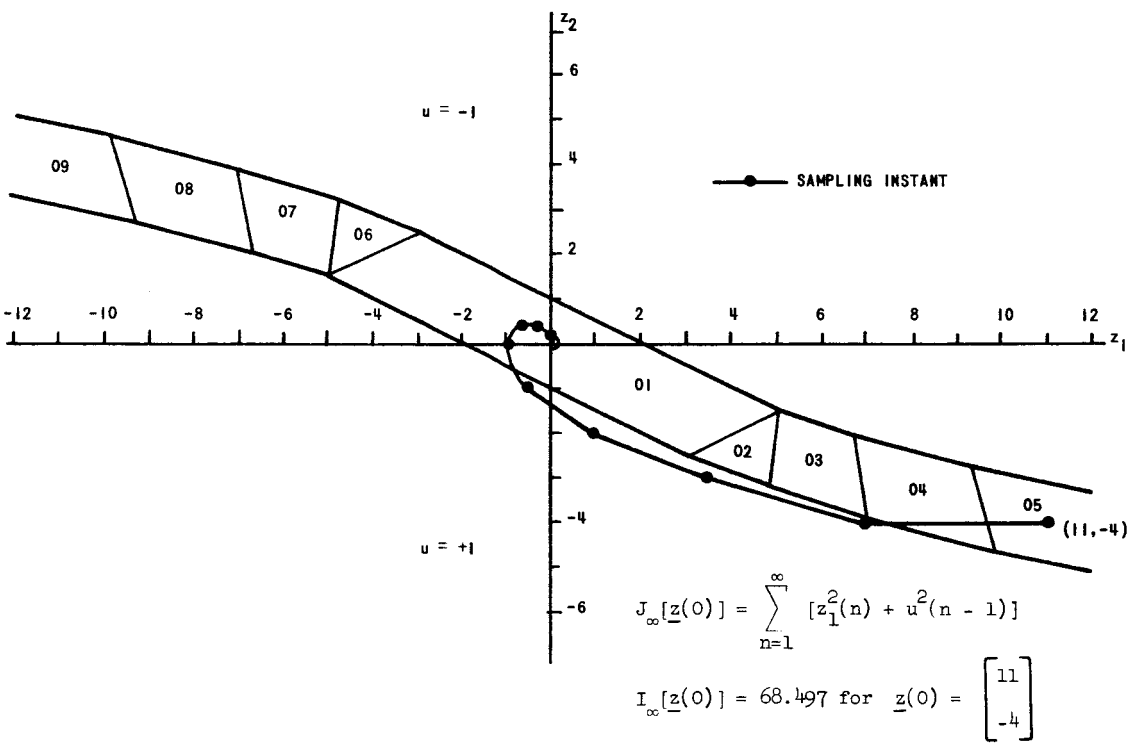

FIG. 6. OPTIMAL CONTROL FOR EXAMPLE A.

Through the middle of Fig. 6 are the regions in which the control is optimally unsaturated, with the region of linear control in the very center. The feedback parameters for these regions are given in Table 1, and the optimal control for each of these regions is given by

$$u(0) = A_{N+1}\underline{z}(0) + B_{N+1} = a_1 z_1(0) + a_2 z_2(0) + b. \qquad (6.5)$$

TABLE 1. FEEDBACK COEFFICIENTS FOR EXAMPLE A

Region No.	A_{N+1}		B_{N+1}
	a_1	a_2	b
01	-0.50000	-1.00000	0.00000
02	-.43902	-1.12195	-0.48780
03	-.35556	-1.13333	-0.92222
04	-.29240	-1.12281	-1.32749
05	-.24658	-1.10959	-1.71918
06	-.43902	-1.12195	0.48780
07	-.35556	-1.13333	0.92222
08	-.29240	-1.12281	1.32749
09	-.24658	-1.10959	1.71918

The alpha-plus regions, where the optimal control is $u = +1$, are shown as the lower part of Fig. 6. Thus the optimal control is determined for every point in the state space shown.

Figure 6 also shows an optimal trajectory starting from initial condition $\underline{z}^T(0) = [11 \quad -4]$. The cost for this initial $\underline{z}(0)$ can be computed either by using Eq. (6.3) or by using

$$I_N[\underline{z}(0)] = \underline{z}^T(0)P_N\underline{z}(0) + \underline{z}^T(0)R_N + R_N^T\underline{z}(0) + C_N, \qquad (6.6)$$

where for the region containing the particular $\underline{z}(0)$ used here,

$$P_N = \begin{bmatrix} 1.56164 & 1.02740 \\ 1.02740 & 1.62329 \end{bmatrix},$$

$$R_N = \begin{bmatrix} -3.94521 \\ -0.25342 \end{bmatrix}, \qquad C_N = 28.74317. \tag{6.7}$$

Calculated either way the optimal cost is 68.497.

[Note that the number N--the number of steps to go--is always infinitely large, since in this problem there are always an infinite number of steps to go. However, since it is necessary to be able to distinguish between the N-stage process and the (N+1)-stage process, the symbol for infinity will not be used to replace N in P_N, R_N, C_N, A_{N+1}, B_{N+1}, or Eq. (6.6).]

In Chapter IV it is stated that the matrix $(\Phi + \Delta A_{N+1})$ is nonsingular when $\Gamma \neq 0$. For the region of linear control in this example (the 01 region in Fig. 6) this matrix is

$$(\Phi + \Delta A_{N+1}) = \begin{bmatrix} 0.75 & 0.50 \\ -0.50 & 0.00 \end{bmatrix} \tag{6.8}$$

The determinant of this matrix is 0.25, and thus the matrix is nonsingular as predicted.

Example A might have been solved using dynamic programming, a general computing method that is able to solve a wide variety of problems, many of which can be solved in no other way. However, the special method used to compute this example needed much less memory storage than dynamic programming would have required. The state-space grid over which this example was computed contained about 30,000 points. A careful use of symmetry might have reduced this to about 20,000 points; even so, dynamic programming would have required at least 60,000 words of storage.

The number of words used by the method of Chapter IV cannot be stated as a function of the size of the state-space grid, since this number depends on whether all unused $\underline{z}(1)$ regions are stored in the fast memory or on tape, on whether $\underline{z}(1)$ is calculated from $\underline{z}(0)$ or vice versa, and on how the regions are stored. In computing this example, only the

boundaries of the regions were stored--a technique that cannot be used in dynamic programming--and all the information about even the largest region, the region of linear control, was stored in less than 200 words. The entire program used only a few thousand words of memory, about one-tenth as many as would have been required by dynamic programming.

B. EXAMPLE B

Consider the same system as used in Example A. The sampling interval is still $\tau = 1$, but the performance index is now

$$J_\infty[\underline{z}(0)] = \sum_{n=1}^{\infty} [z_1^2(n) + z_2^2(n)]. \tag{6.9}$$

Thus

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \Gamma = 0. \tag{6.10}$$

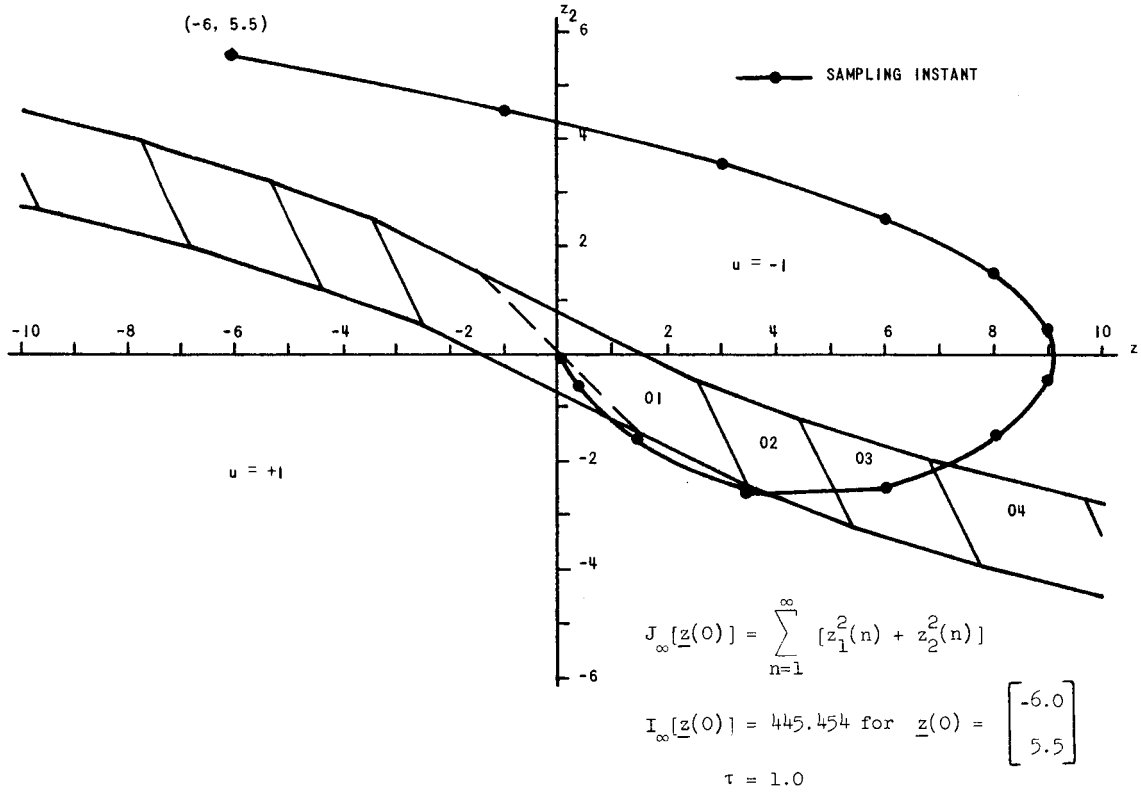

FIG. 7. OPTIMAL CONTROL FOR EXAMPLE B.

The optimal control is shown in Fig. 7 along with an optimal trajectory. An extended picture of the optimal control is given in Fig. 8. Because the bounds on the control are symmetrical, the regions

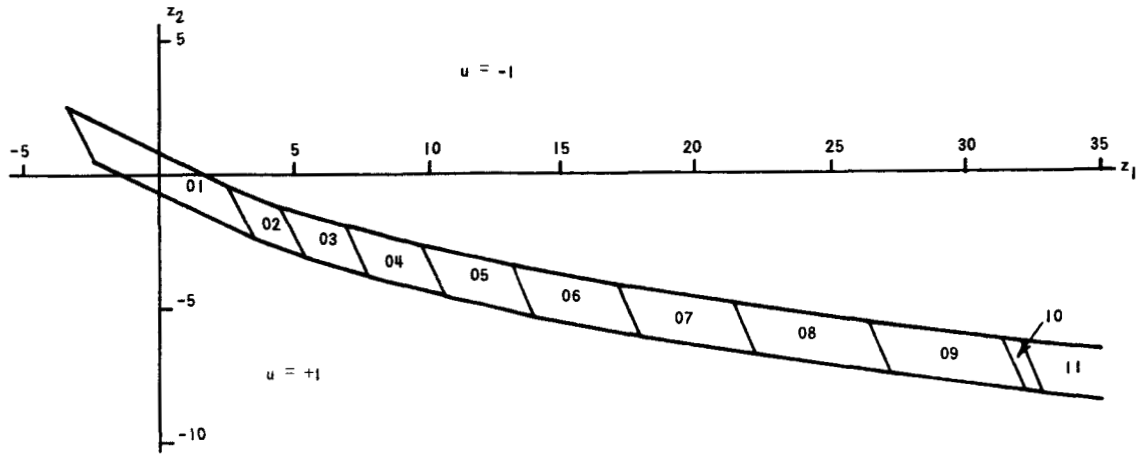

FIG. 8. EXTENDED REGION OF OPTIMAL CONTROL FOR EXAMPLE B.

are symmetric with respect to the origin, and thus only half the unsaturated regions are shown in Fig. 8. The feedback parameters A_{N+1} of symmetric regions have the same value and sign, while the B_{N+1} have the opposite sign. The feedback parameters for both Figs. 7 and 8 are given in Table 2.

The optimal trajectory shown in Fig. 7 begins at $\underline{z}^T(0) = [-6.0 \quad 5.5]$. The cost associated with this initial condition can be computed by using either Eq. (6.6) or (6.9). In either case the cost is 445.454, where for the alpha-minus region containing $\underline{z}(0)$,

$$P_N = \begin{bmatrix} 9.28916 & 46.95783 \\ 46.95783 & 305.14152 \end{bmatrix},$$

$$R_N = \begin{bmatrix} -149.32529 \\ -1086.2649 \end{bmatrix}, \qquad C_N = 4136.7404. \tag{6.11}$$

TABLE 2. FEEDBACK COEFFICIENTS FOR EXAMPLE B

Region No.	A_{N+1}		B_{N+1}
	a_1	a_2	b
01	-0.66667	-1.33333	0.00000
02	-.47058	-1.23529	-0.44118
03	-.36145	-1.18072	-0.85542
04	-.29268	-1.14634	-1.25610
05	-.24561	-1.12281	-1.64912
06	-.21145	-1.10573	-2.03744
07	-.18557	-1.09278	-2.42268
08	-.16528	-1.08264	-2.80578
09	-.14898	-1.07449	-3.18735
10	-.14792	-1.07396	-3.21698
11	-.13559	-1.06779	-3.56779

It is proved in Chapter IV that if $\Gamma = 0$ the matrix $(\Phi + \Delta A_{N+1})$ is singular. For this example the matrix for the region of linear control is

$$(\Phi + \Delta A_{N+1}) = \begin{bmatrix} 2/3 & 1/3 \\ -2/3 & -1/3 \end{bmatrix} \tag{6.12}$$

which is certainly singular. Direct calculation shows that for any of the A_{N+1} calculated in this example this matrix is singular.

Computing the minimum cost from Eq. (6.9) requires summing an infinite series. This is particularly easy in this case where the matrix (6.12) is singular. If $\underline{z}(0)$ is in the region of linear control, then the $\underline{z}(n)$ $(n = 1, 2, \ldots)$ always lie on a line through the origin, in this case the line with slope -1 shown dashed in Fig. 7. Since $z_1(n) = -z_2(n)$, using (6.12) with (6.1) shows that each $\underline{z}(n + 1)$ is given by the geometrical progression

$$\underline{z}(n + 1) = \frac{\underline{z}(n)}{3} \qquad n = 1, 2, \ldots \tag{6.13}$$

Thus the optimal cost is given by

$$I_\infty[\underline{z}(0)] = \sum_{n=1}^{\infty} \frac{z_1^2(1) + z_2^2(1)}{9^{n-1}} = 2z_1^2(1) \sum_{n=1}^{\infty} \frac{1}{9^{n-1}} = \frac{9z_1^2(1)}{4}. \quad (6.14)$$

C. EXAMPLE C

This example shows the effect of increasing the sampling rate. The system is the same as in Example B, but the sampling interval $\tau = 0.1$ is one-tenth as long.

As shown in Fig. 9, the band of regions of unsaturated control is much narrower than in Fig. 7 of Example B. Since the regions are much smaller, there are many, many more of them. There are over 50 regions of unsaturated control on each side of the region of linear control in the state space shown in Fig. 9. The boundaries separating these regions are not shown because they are so close together.

The optimal control for the region of linear control is given by

$$u = (-9.52382)z_1 + (-10.47619)z_2. \quad (6.15)$$

The A_{N+1} for the region of linear control was calculated, beginning with $P_0 = 0$, by the iteration method discussed in Chapter IV. A_{N+1} in Example B took eight iterations to converge to six significant figures; in Example C it took about 80 iterations to converge to the same number of figures.

D. EXAMPLE D

This example, perhaps the most interesting presented here, considers the artificial earth satellite, including the external force due to the gravity gradient, discussed in Chapter II. All parameters are normalized and the sampling interval is arbitrarily and somewhat unrealistically set at $\tau = 1$. Thus the state-vector-transition equation is

$$\underline{z}(n+1) = \begin{bmatrix} 0.54030 & 0.84147 \\ -0.84147 & 0.54030 \end{bmatrix} \underline{z}(n) + \begin{bmatrix} 0.45970 \\ 0.84147 \end{bmatrix} u(n) \quad (6.16)$$

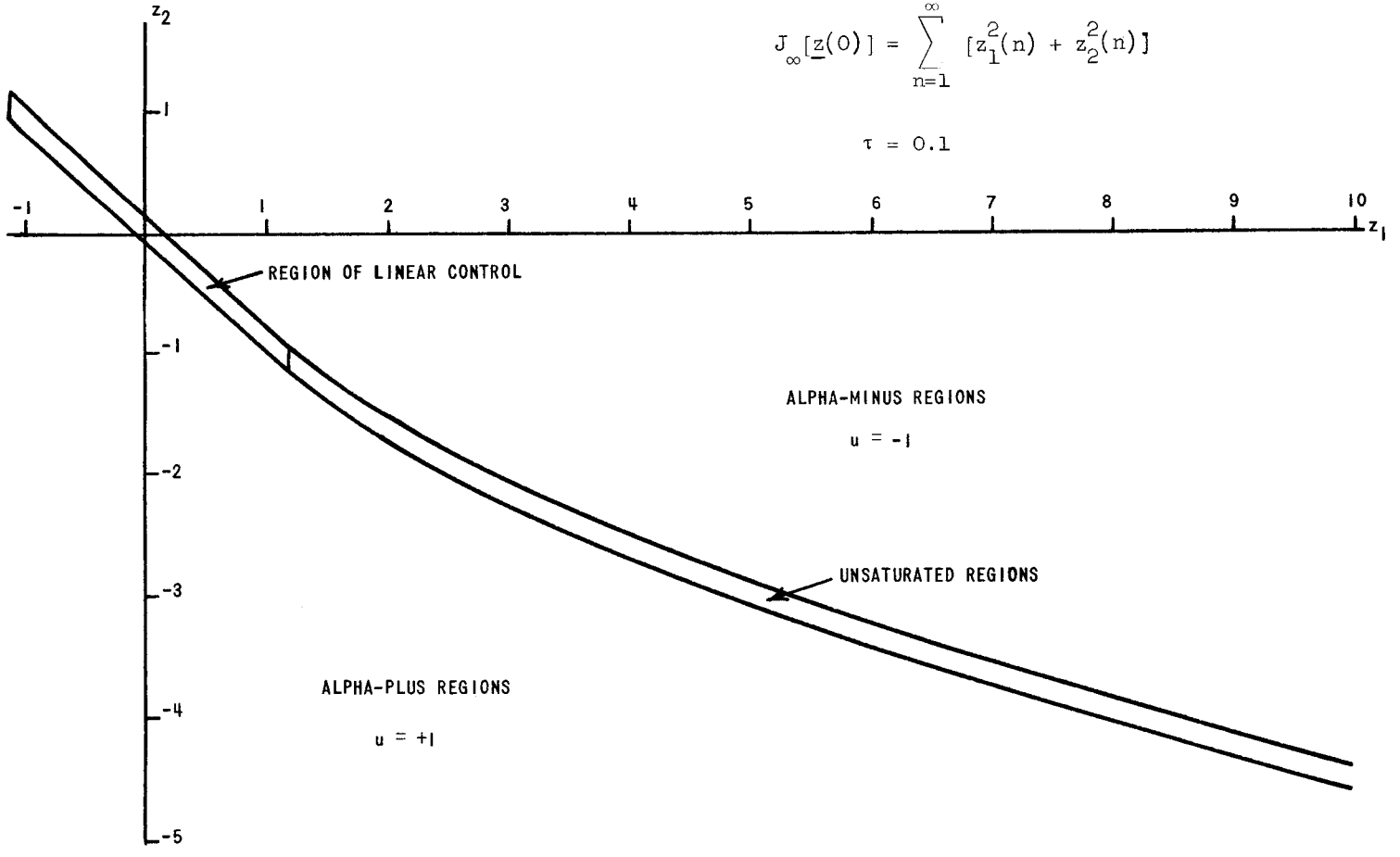

FIG. 9. OPTIMAL CONTROL FOR EXAMPLE C.

where the control is limited by Eq. (6.2). This system is shown in Fig. 10.

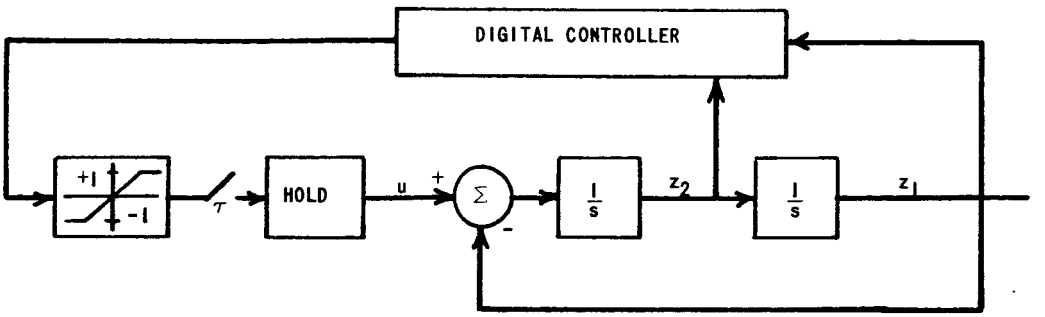

FIG. 10. BLOCK DIAGRAM OF SYSTEM FOR EXAMPLE D.

It is desired to make the attitude integral-squared-error a minimum, and as an approximation the attitude sum-squared-error will be minimized. (This approximation will unfortunately cause a phenomenon known as inter-sample ripple, as will be shown later.) There is no cost on the control, thus the performance index is

$$J_\infty[\underline{z}(0)] = \sum_{n=1}^{\infty} z_1^2(n) \tag{6.17}$$

Therefore

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \Gamma = 0. \tag{6.18}$$

For the purpose of comparison, a simple nonoptimal system is shown in Fig. 11. This system uses the optimal feedback gains a_1 and a_2 of the region of linear control, and thus is optimal for initial conditions close to the origin in state space. If the total output of these feedback blocks causes saturation of the input, then this is simply allowed to happen. Thus Fig. 11 shows a system that is extremely simple to build and is optimal for small initial conditions.

The optimal control regions are shown in Fig. 12; the optimal feedback coefficients for the unsaturated regions are given in Table 3.

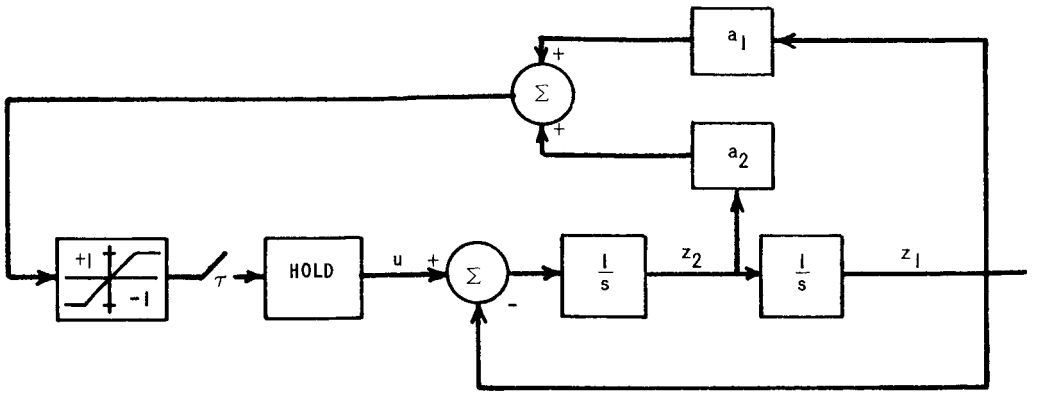

FIG. 11. NONOPTIMAL SYSTEM OF EXAMPLE D.

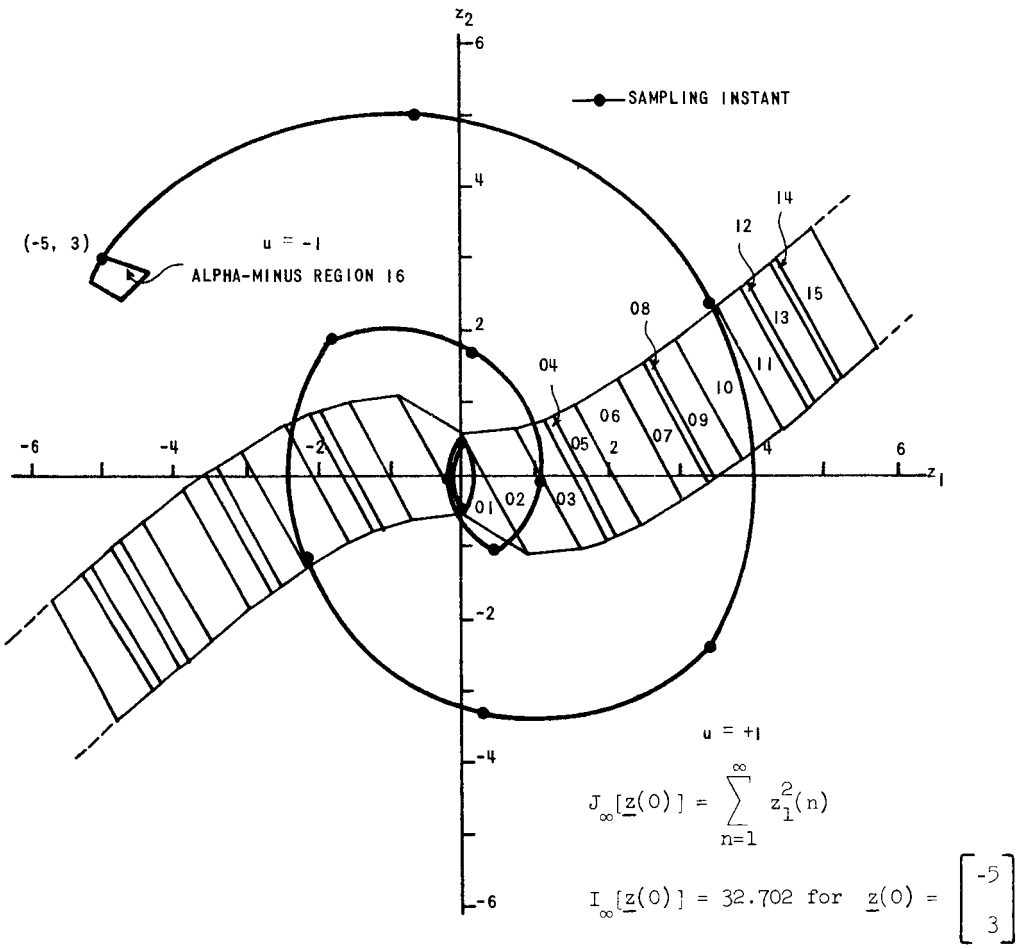

FIG. 12. OPTIMAL CONTROL FOR EXAMPLE D.

TABLE 3. FEEDBACK COEFFICIENTS FOR EXAMPLE D

Region No.	A_{N+1}		B_{N+1}
	a_1	a_2	b
01	-1.17534	-1.83049	0.00000
02	0.13289	-1.11580	-0.39044
03	0.35351	-0.99526	-0.62492
04	0.31743	-1.01498	-0.56883
05	0.49316	-0.91898	-0.86046
06	0.53957	-0.89363	-0.95780
07	0.55971	-0.88262	-1.01570
08	0.66825	-0.82333	-1.38211
09	0.60572	-0.85749	-1.16421
10	0.64942	-0.83362	-1.33612
11	0.66819	-0.82336	-1.42508
12	0.76747	-0.76912	-1.94701
13	0.71067	-0.80016	-1.64189
14	0.75047	-0.77841	-1.87443
15	0.63394	-0.84207	-1.17957
16	0.49844	-0.91610	1.96661

These feedback coefficients are used in Eq. (6.5) to determine the optimal control.

An optimal trajectory from initial condition $\underline{z}^T(0) = [-5.0 \quad 3.0]$ is shown in Fig. 12. This initial condition is in alpha-minus region 16 as shown in the figure. When the state vector reaches the region of linear control, it enters a limit cycle rather than going to the origin. The value of $z_1(n)$ at the sampling instants is zero in this limit cycle, thus no cost is charged to the performance index (6.17). However, the attitude error $z_1(t)$ is zero only at the sampling instants, and thus there exists a phenomenon known as intersample ripple [Ref. 15]. This ripple can be eliminated during the design of the system by adding a charge on either $z_2(n)$ or $u(n - 1)$.

The cost of the trajectory shown in Fig. 12, as determined either by (6.17) or by (6.6), is 32.702. This cost is most easily determined by (6.6), where

$$P_N = \begin{bmatrix} 2.947579 & -0.831075 \\ -0.831075 & 4.033005 \end{bmatrix},$$

$$R_N = \begin{bmatrix} 9.389736 \\ -9.400775 \end{bmatrix}, \qquad C_N = 48.085094 \tag{6.19}$$

for the region containing $\underline{z}(0)$--alpha-minus region 16 in Fig. 12.

The control for the nonoptimal system of Fig. 11 is given in Fig. 13. Between the two parallel lines the control is given by

$$u = (-1.175343)z_1 + (-1.830488)z_2 \tag{6.20}$$

where the two lines are determined by setting $u = \pm 1$ in (6.20). The trajectory shown in Fig. 13 from initial condition $\underline{z}^T(0) = [-5.0 \quad 3.0]$ has a cost determined from Eq. (6.17) of 73.154, an increase of 124 percent over the optimal system. Thus, though considerably more complicated to mechanize, the optimal system is a substantial improvement over the simple system of Fig. 11.

E. THE SYNTHESIS

The optimal design of a system is often used only as a standard of comparison for the system that is actually built. However, if the truly optimal system is to be synthesized, the feedback coefficients of the unsaturated regions can be stored in a special-purpose digital computer. The computer takes the value of the state vector at the sampling instant and decides whether the control u is optimally α^+, α^-, or unsaturated. If u is optimally unsaturated, the computer determines which unsaturated region the state vector is in, and computes the control using Eq. (6.5).

Approximations to the optimal system can be made with varying degrees of accuracy. For example, some of the unsaturated regions can be combined into one region with little deviation from the optimal cost; or all

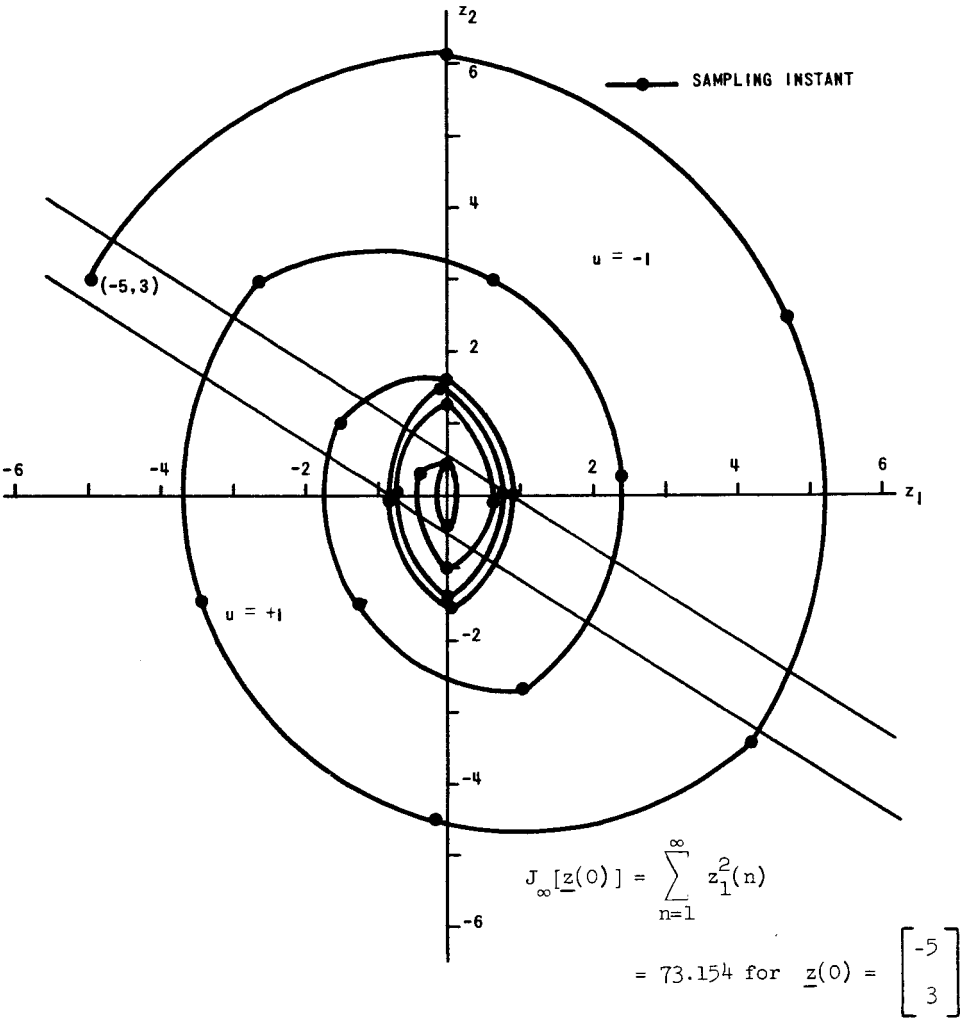

FIG. 13. CONTROL FOR SYSTEM OF FIG. 11.

of the unsaturated regions can be eliminated, using their location as a guide to the placement of a piecewise linear switching curve. Around the origin in state space, however, the control must be linear if the system is to return to equilibrium.

REFERENCES

1. R. E. Kalman and J. E. Bertram, "A Unified Approach to the Theory of Sampling Systems," J. Franklin Inst., 267, 5, May, 1959, pp. 405-436.

2. R. E. Kalman and R. W. Koepcke, "Optimal Synthesis of Linear Sampling Control Systems Using Generalized Performance Indexes," ASME Trans., 80, 1958, pp. 1820-1826.

3. E. W. Henry, "Logical Scheduling of a Multiplexed Digital Controller," TR No. 2106-1, Contract Nonr 225(38), Stanford Electronics Laboratories, Stanford, Calif., Jul 1960.

4. F. Kurzweil, Jr., "The Analysis and Synthesis of Nonlinear Continuous and Sampled-Data Systems Involving Saturation," TR No. 2101-1, Contract Nonr 225(24), Stanford Electronics Laboratories, Stanford, Calif., 30 Nov 1959.

5. C. A. Desoer, and J. Wing, "A Minimal Time Discrete System," IRE Tran. on Automatic Control, AC-6, 2, May 1961, pp. 111-125.

6. R. E. Kalman, "Optimal Nonlinear Control of Saturating Systems by Intermittent Action," IRE Wescon Conv. Rec., Part 4, Aug 1957, pp. 130-135.

7. C. W. Merriam III, "A Class of Optimum Control Systems," J. Franklin Inst., 267, 4, Apr 1959, pp. 267-281.

8. C. W. Merriam III, "An Optimization Theory for Feedback Control System Design," Information and Control, Mar 1960, pp. 32-59.

9. R. Bellman, Dynamic Programming, Princeton University Press, Princeton, N. J., 1957.

10. R. E. Bellman and S. E. Dreyfus, Applied Dynamic Programming, Princeton University Press, Princeton, N. J., 1962.

11. R. E. Kalman, "On the General Theory of Control Systems," Proc. First International Congress on Automatic Control (Moscow, 1960), Butterworth's, 1961.

12. T. L. Gunckel, II, "Optimum Design of Sampled-Data Systems with Random Parameters," TR No. 2102-2, Contract Nonr 225(24), Stanford Electronics Laboratories, Stanford, Calif., 24 Apr 1961.

13. H. E. Rauch, "Linear Estimation of Sampled Stochastic Processes with Random Parameters," Rept. SEL-62-058 (TR No. 2108-1), Stanford Electronics Laboratories, Apr 1962.

14. R. H. Cannon, Jr., "Some Basic Response Relations for Reaction-Wheel Attitude Control," ARS Jour., 32, 1, Jan 1962, pp. 61-74.

15. J. R. Ragazzini and G. F. Franklin, Sampled-Data Control Systems, McGraw-Hill Book Co., Inc., New York, 1958.

16. R. E. Roberson, "Attitude Control of a Satellite Vehicle--An Outline of the Problems," Proc. VIIIth International Astronautical Congress, Barcelona, 1957, Wien, Springer-Verlag, 1958, pp. 317-339.

17. R. Bellman, Introduction to Matrix Analysis, McGraw-Hill Book Co., Inc., New York, 1960.